An invitation to quantum filtering and smoothing theory based on two inner products (Building Foundations for Quantum Statistical Modeling)

Ohki, Kentaro

数理解析研究所講究録
第2018巻 2017年 18-44

An invitation to quantum filtering and smoothing theory based on two inner products

By

Kentaro OHKI*

Abstract

The purpose of this paper is to introduce the quantum filtering and a smoothing theory for Markovian open quantum dynamical systems briefly. The filtering and smoothing theory for classical systems are well developed and show their performance used in practice, and the quantum filtering theory has also developed from 1980s based on the quantum version of the classical conditional expectation. However, the quantum smoothing theory has not developed since the quantum conditional expectation is not defined as well. We developed a quantum smoothing theory based on two quantum inner products and show it with the filtering theory. The weak value is naturally defined as the minimum mean square estimate in our frame work and the quantum smoother, the dynamical estimator of weak value, is derived.

§ 1. Introduction

State estimation problems of dynamical systems driven by noisy disturbances often arise in practice, and many kinds of estimators are proposed to extract the exact signal from noisy observations [30, 49, 50, 56, 60]. A state estimation problem is a problem that reconstruction of the objective, such as parameters of a probability density function and physical quantities, from measured data (Fig. 1). One of the central notions of state estimations is the so-called *minimum mean squares estimation* and it is widely used because it is intuitive and easy to obtain the optimal solution by Hilbert space theory [6, 38, 40].

Despite the first study of the minimum mean squares estimation problems was found in 1700s [51], it has been studied and developed so far; for examples, a relation between estimation errors and mutual information is found and developed in [17, 28, 29, 57],

Received April 20, 201x. Revised September 11, 201x.
2010 Mathematics Subject Classification(s):
Key Words: *Quantum filtering theory, Quantum smoothing theory*
*Graduate School of informatics, Kyoto University, Kyoto 606-8501, Japan.
e-mail: `ohki@i.kyoto-u.ac.jp`

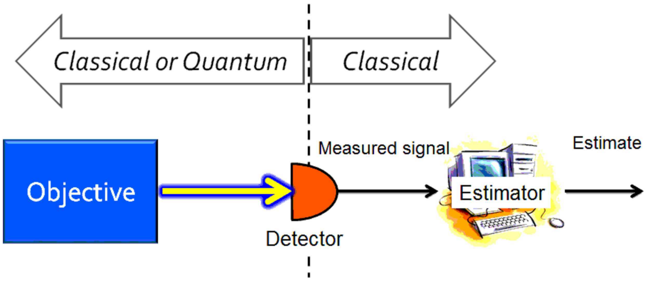

Figure 1. Abstract setting of estimation problems

and new estimators for dynamical systems described on certain manifolds, quantum systems are derived [13, 14, 23]. The minimum mean square estimator is given by the conditional expectation and we obtain a recursive estimator for hidden random variables [38, 40]. First of all, let us review the dynamical minimum mean square estimator without rigorous descriptions. Consider a system driven by Wiener noise with noisy measurement:

$$\begin{aligned} dX_t =& f(X_t)dt + g(X_t)dW_t, \\ dY_t =& h(X_t)dt + dV_t, \end{aligned}$$

where all signals are one-dimensional random variables at any time $t \geq 0$. W and V are mutually independent one-dimensional standard Wiener processes. The functions f, g and h satisfy certain suitable conditions. Let $\mathcal{C}_t$ be a set of $\mathcal{F}_t := \sigma(\{Y_s \mid 0 \leq s \leq t\})$–measurable real–valued functions and P be a probability measure of the whole signals.

Problem 1.1.
Find a function $Q^{\rm opt} \in \mathcal{C}_t$ as a solution of the following minimization problem

$$\min_{Q \in \mathcal{C}_t} \mathbb{E}_P \left[|X_\tau - Q|^2 \right],$$

where $\mathbb{E}_P$ is the expectation with respect to the probability measure P.

Problem 1.1 is called the prediction problem if $\tau > t$, the filtering problem if $\tau = t$, and the smoothing problem if $\tau < t$, respectively. Whether t is larger than τ or not, the solution of Problem 1.1 is given by the $Q^{\rm opt} \in \mathcal{C}_t$ satisfying

$$\mathbb{E}_P \left[(X_\tau - Q^{\rm opt}) Z \right] = 0, \quad \forall Z \in \mathcal{C}_t.$$

This implies that the estimation error and $\mathcal{C}_t$ are mutually orthogonal under the probability measure P. The orthogonality defines the conditional expectation [11], and $Q^{\rm opt}$

is usually denoted by $\mathbb{E}_P\left[X_\tau|\mathcal{F}_t\right]$. The filtering and fixed point smoothing equations are then described of following forms

$$d\pi_t(X) = \pi_t(f(X))dt + \mathbb{E}_P\left[(X_t - \pi_t(X))\left(h(X_t) - \pi_t(h(X))\right)\middle|\mathcal{F}_t\right](dY_t - \pi_t(h(X))dt),$$
$$d\pi_{\tau|t}(X) = \mathbb{E}_P\left[(X_\tau - \pi_{\tau|t}(X))\left(h(X_t) - \pi_t(h(X))\right)\middle|\mathcal{F}_t\right](dY_t - \pi_t(h(X))dt).$$

As in classical statistics, estimation problems have also been arisen in quantum physics [32, 33, 34]. Since current technologies give us many experiments of quantum systems [20, 39, 45, 46, 47, 48], the necessity of the statistical procedure for quantum systems has been increasing. According to the quantum theory, every quantum physical quantity, even a measured signal, is described by a self–adjoint operator on certain Hilbert space $\mathcal{H}$. Then the estimation problem is the estimation of the self-adjoint operators rather than that of the random variables. Consider the following quantum system without details:

$$\begin{aligned} d\hat{X}_t =& \hat{\mathcal{L}}(\hat{X}_t)dt + [\hat{L}_t^*, \hat{X}_t]_- d\hat{A}_t + [\hat{X}_t, \hat{L}_t]_- d\hat{A}_t^*, \\ d\hat{Y}_t =& \left(\hat{L}_t + \hat{L}_t^*\right) dt + d\hat{A}_t + d\hat{A}_t^*, \end{aligned}$$

where $\hat{\mathcal{L}}$ is a linear map, which is called the Lindblad operator , $\hat{L}_t$ is a coupling operator with quantum noisy environment, and $\hat{A}_t$ is a quantum noise. $\hat{X}_t$ is the quantum physical quantity estimand-to-be and $\hat{Y}_t$ is the measured signal. Unlike the classical probability theory, linear operators on Hilbert space are used to describe the quantum random variables. Since the measurement data is classical, we use certain commutative operator algebra $\mathcal{Y}_t$, called a commutative *von Neumann algebra*, which implies $\mathcal{F}_t$–measurable functions. A quantum counter part of real-valued measurable function is then a self–adjoint operator in $\mathcal{Y}_t$. From the above situation, consider following problem.

Problem 1.2. *Find an operator $\hat{Q}^{\rm opt} \in \mathcal{Y}_t$ as a solution of the following minimization problem*

$$\min_{\hat{Q}\in\mathcal{Y}_t} \mathbb{P}_{\hat{\rho}}\left[\left|\hat{X}_\tau - \hat{Q}\right|^2\right],$$

where $|\hat{A}|^2 := \hat{A}^*\hat{A}$ and $\mathbb{P}_{\hat{\rho}}$ is the quantum expectation with respect to the density operator $\hat{\rho}$.

As we discuss details in the following sections, there exists the optimum solution $\hat{Q}^{\rm opt}$ even if the object is described in quantum mechanics. For the filtering problem, the quantum filter is obtained by dynamical representation of a quantum version of conditional expectation [13, 14]. The optimal solution $\hat{Q}^{\mathrm{opt}}$ of the quantum filtering problem satisfies the following orthogonality,

$$\mathbb{P}_{\hat{\rho}}\left[\hat{Z}\left(\hat{X}_\tau - \hat{Q}^{\mathrm{opt}}\right)\right] = 0, \quad \forall \hat{Z} \in \mathcal{Y}_t, \quad \tau = t.$$

Usually the $\hat{Q}^{\mathrm{opt}}$ is denoted by $\mathbb{P}_{\hat{\rho}}[\hat{X}_\tau|\mathcal{Y}_t]$. Then the recursive filtering equation is

$$\begin{aligned} d\hat{\pi}_t(\hat{X}) =& \hat{\pi}_t\left(\hat{\mathcal{L}}(\hat{X})\right) dt \\ &+ \mathbb{P}_{\hat{\rho}}\left[(\hat{L}_t - \hat{\pi}_t(\hat{L}))^*\hat{X}_t + \hat{X}_t(\hat{L}_t - \hat{\pi}_t(\hat{L}))\Big|\mathcal{Y}_t\right](d\hat{Y}_t - \hat{\pi}_t(\hat{L} + \hat{L}^*)dt), \end{aligned}$$

where $\hat{\pi}_t(\hat{X}) := \mathbb{P}_{\hat{\rho}}[\hat{X}_t|\mathcal{Y}_t]$ is the quantum conditional expectation. The solutions of quantum prediction and filtering problems are obtained by the quantum conditional expectation [8, 13, 14] and the quantum conditional expectation is well-defined if we consider the indirect measurement [61]. The key notion why the quantum conditional expectation is well-defined is the *commutativity* between measurement records and physical quantities to be estimated, and it ensures a certain orthogonal condition between estimation error and measurement records. The quantum filtering theory also shed a light on the measurement–based feedback control theory for quantum systems [3, 9, 10, 21, 37, 61, 64].

However, the solution of the general quantum smoothing problems is not described by the quantum conditional expectation. In contrast to the classical random variables, quantum random variables do not have the commutativity with respect to multiplication and the past physical quantities does not commute with the measurement records in general. This makes ones impossible to define the quantum conditional expectation, therefore, the general smoothing theory must not be based on the quantum conditional expectation. The previous work on quantum smoothing problems, Yanagisawa found that the quantum systems which smoothing problem is described by the quantum conditional expectation [62]. This work opened the door of the quantum smoothing problems and several researches have tackled with the problems [25, 53, 54, 63]. For example, Tsang also gave a smoothing method for a quantum phase estimation problem [53], which is based on the time-symmetric approach proposed by Aharonov et al. [2]. Yonezawa et al. [63] gave another approach for the quantum optical-phase estimation and showed experimental results of the estimation problem. Since these estimation problems are the estimation of the classical parameter in the quantum system, they end up solving the classical smoothing problem. Recently Gammelmark et al. derived a new past state estimation scheme [25] based on weak values [1, 22]. Furthermore, the author of this paper proposed a new smoothing theory based on two inner products [42, 43], which is the main topic of this paper.

In this paper, we show the optimal solution $\hat{Q}^{\mathrm{opt}}$ of Problem 1.2 is composed of two operators $\hat{Q}^{\mathrm{opt}} = \hat{Q}^+ + \hat{Q}^-$, where $\hat{Q}^+$ and $\hat{Q}^-$ satisfying *the symmetric orthogonal* *condition* and *the skew symmetric condition* under a quantum state $\hat{\rho}$

$$\mathbb{P}_{\hat{\rho}}\left[\hat{Z}\left(\hat{X}_\tau - \hat{Q}^+\right) + \left(\hat{X}_\tau - \hat{Q}^+\right)\hat{Z}\right] = 0,$$
$$\mathbb{P}_{\hat{\rho}}\left[\hat{Z}\hat{X}_\tau - \hat{X}_\tau\hat{Z}\right] = 2\mathbb{P}_{\hat{\rho}}\left[\hat{Q}^-\hat{Z}\right], \quad \forall\hat{Z} \in \mathcal{Y}_t,$$

respectively. We also show that $\hat{Q}^+$ is the best approximation in the sense of symmetric inner product. The optimal solution $\hat{Q}^{\rm opt}$ coincides with the "weak value" of $\hat{X}$, which is well-known in physics literatures [1, 18, 19, 22, 24]. Furthermore, we propose a new framework for quantum smoothing theory based on *the symmetric orthogonal condition.* This is a natural extension of the quantum conditional expectation and gives a recursive minimum mean square estimation for past quantum physical quantities. Recently, Amini et al. derived the linear minimum mean squares estimator for linear quantum systems [5] and the estimator is realized in quantum systems. Their estimator is interesting but it is impossible to estimate past quantum states because the implemented estimator is causal. Our proposal estimator is for general Markovian quantum systems and realized in classical systems. It is possible to implement non-causal estimator in practice, so it is possible to estimate the past quantum states in principle.

The rest of this paper is organized as follows. In Section 2, we introduce some foundations of quantum theory and quantum statistics. Especially, we show the key idea of the main result of this paper in finite dimensional quantum system. The concept of this paper is described in this section. In Section 3, we introduce the symmetric orthogonality and asymmetric condition and show that the operators satisfying these conditions are the real part and imaginary part of the minimum mean square estimation. The quantum dynamical system considered in this paper and its filter are shown in Section 4. We develop a new quantum smoother in Section 5 and conclude this paper in Section 6.

Notation

$\mathbb{R}$ and $\mathbb{C}$ are real numbers and complex numbers, respectively, and $\mathrm{i} := \sqrt{-1}$. $\mathcal{H}$ is a complex Hilbert space and we also denote $\mathcal{H}_X$ if it is the Hilbert space of the system X. Any linear operator on a Hilbert space $\mathcal{H}$ is denoted by hat, e.g., $\hat{X}$. When positive operators $\hat{X}$ and $\hat{Y}$ satisfy $\hat{X} = \hat{Y}^2$, we denote $\hat{Y} = \sqrt{\hat{X}}$. The absolute value of operator is defined by $|\hat{X}| := \sqrt{\hat{X}^*\hat{X}}$. $\mathcal{L}(\mathcal{H})$ is the set of all linear bounded operators on the Hilbert space $\mathcal{H}$. $\hat{X} \geq 0$ means that $\hat{X} \in \mathcal{L}(\mathcal{H})$ is a positive operator and $\hat{X}^*$ implies the conjugate operator of $\hat{X}$. $\mathrm{Tr}[\bullet] : \mathcal{L}(\mathcal{H}) \to \mathbb{C}$ is the trace on the linear bounded operators. $\mathcal{S}(\mathcal{H}) := \{\hat{\rho} \in \mathcal{L}(\mathcal{H}) \mid \hat{\rho} \geq 0, \mathrm{Tr}[\hat{\rho}] = 1\}$ is a set of density operators. $\hat{1}_{\mathcal{H}}$ is the identity operator on $\mathcal{H}$ and we sometimes omit its subscript. Denote $[\hat{X}, \hat{Y}]_\pm := \hat{X}\hat{Y} \pm \hat{Y}\hat{X}$, $\forall\hat{X}, \hat{Y} \in \mathcal{L}(\mathcal{H})$. $\otimes$ represents the Kronecker product for matrices and the tensor product for operators, Hilbert spaces, or the sets of linear operators.

§ 2. Basics of quantum theory and estimation

§ 2.1. Basics of quantum probability theory

In this section, we briefly review the quantum theory (for details, see, e.g., [16, 33, 41, 44, 61].) Quantum physics is described by a generalized probability theory, called *quantum probability theory* or *noncommutative probability theory.* It is essentially a probability theory that consists of matrices. In this section, we introduce the quantum probability theory from the elementary probability theory [44, 33, 35].

Consider a set of outcomes $\{x_1, \cdots, x_n\}$ of the random variable X and an n-dimensional variable $x = (x_i) \in \mathbb{R}^n$ with probability vector $p = (p_i)$; $p_i \geq 0$, $\sum_{i=1}^n p_i = 1$. Then the expectation of the random variable x under probability vector p can be described as

$$\begin{aligned}\sum_{i=1}^n p_i x_i =&\mathrm{Tr}\left[\begin{bmatrix} p_1 & & & \\ & p_2 & & \\ & & \ddots & \\ & & & p_n \end{bmatrix}\begin{bmatrix} x_1 & & & \\ & x_2 & & \\ & & \ddots & \\ & & & x_n \end{bmatrix}\right] \\ =&\mathrm{Tr}\left[\begin{bmatrix} p_1 & * & * & * \\ * & p_2 & * & * \\ * & * & \ddots & * \\ * & * & * & p_n \end{bmatrix}\begin{bmatrix} x_1 & & & \\ & x_2 & & \\ & & \ddots & \\ & & & x_n \end{bmatrix}\right] = \mathrm{Tr}[\hat{\rho}\hat{X}].\end{aligned}$$

From the cyclic property of the trace, $\mathrm{Tr}[\hat{\rho}\hat{X}] = \mathrm{Tr}[\hat{V}\hat{\rho}\hat{V}^*\hat{V}\hat{X}\hat{V}^*]$ for any unitary matrix $\hat{V} \in \mathbb{C}^{n\times n}$, so $\hat{\rho}$ and $\hat{X}$ can be represented by Hermitian matrices. The $\hat{\rho} \in \mathbb{C}^{n\times n}$ is an extension of probability vector for matrices and the Hermitian matrix $\hat{X} \in \mathbb{C}^{n\times n}$ is a matrix-version random variables. $\hat{\rho}$ is called a *density matrix* if $\hat{\rho} \geq 0$ and $\mathrm{Tr}[\hat{\rho}] = 1$. In this paper, we call a Hermitian matrix *a quantum random variable* or *a quantum physical quantity* whether it represents a real physical quantity or not. An outcome of measurement of a quantum random variable is one of its eigenvalue with certain probability determined by $\hat{\rho}$ and the measurement setup [33, 34, 35]. A probability theory with Hermitian matrices and density matrices $\{\hat{\rho} \in \mathbb{C}^{n\times n} \mid \hat{\rho} \geq 0,\ \mathrm{Tr}[\hat{\rho}] = 1\}$ is called *quantum probability theory,* which describes the statistical structure and probabilistic nature of quantum physics. We can find a $*$–isomorphism ι for a quantum random variable $\hat{X} = \hat{X}^* \in \mathbb{C}^{n\times n}$ such that $\iota(\hat{X}) = x \in \mathbb{R}^n$, where $x = (x_i)$ is a vector which elements are eigenvalues of $\hat{X}$. The corresponding classical random variable $X : \Omega = \{1, 2, \cdots, n\} \to \mathbb{R}$ is defined by $X(i) = x_i$.

In general a quantum system is described by a suitably defined Hilbert space $\mathcal{H}$. Any physical quantity of a quantum system is denoted by self-adjoint operator $\hat{X}$ on

$\mathcal{H}$. Although most of physical quantities are described by unbounded self-adjoint operators in practice, we only consider linear bounded operators except quantum noise operators introduced in the later section. We denote a set of linear bounded operators on $\mathcal{H}$ by $\mathcal{L}(\mathcal{H})$. Since we only use operators with any finite moments under the given state defined below, this is not strict constraint. The observation of any quantum physical quantity is a randomly chosen number from the spectrum of the corresponding self-adjoint operator. Random outcomes of all bounded operators make the quantum statistics and the quantum expectation $\mathbb{P}_{\hat{\rho}}$ is defined as $\mathbb{P}_{\hat{\rho}}[\hat{X}] = \mathrm{Tr}[\hat{\rho}\hat{X}]$, $\hat{\rho} \in \mathcal{S}(\mathcal{H})$. In contrast to classical probability space, the quantum version of the set of the measurable functions is defined as a von Neumann Algebra [52]. Roughly speaking, it is an algebra generated by projection operators with algebraic operations [14]. Let $\mathcal{N} \subseteq \mathcal{L}(\mathcal{H})$ be von Neumann subalgebra. A pair $(\mathcal{N}, \mathbb{P}_{\hat{\rho}})$ is called *the quantum probability space*. For a given quantum probability space $(\mathcal{N}, \mathbb{P}_{\hat{\rho}})$, a subalgebra $\mathcal{N}_{\hat{\rho}} := \{\hat{X} \in \mathcal{N} \mid \mathbb{P}_{\hat{\rho}}[\hat{X}^*\hat{X}] = 0\}$ of $\mathcal{N}$ is a quantum version of the measure zero set with respect to $\mathbb{P}_{\hat{\rho}}$, called the left kernel of $\mathbb{P}_{\hat{\rho}}$. The left kernel $\mathcal{N}_{\hat{\rho}}$ is not empty since it always includes 0. Moreover, $\mathcal{N}_{\hat{\rho}}$ is a left ideal and satisfies

$$\mathbb{P}_{\hat{\rho}}\left[(\hat{X} + \hat{Z}_1)^*(\hat{Y} + \hat{Z}_2)\right] = \mathbb{P}_{\hat{\rho}}\left[\hat{X}^*\hat{Y}\right]$$

for any $\hat{X}, \hat{Y} \in \mathcal{N}$ and $\hat{Z}_1, \hat{Z}_2 \in \mathcal{N}_{\hat{\rho}}$ (see Lemma 9.6 of Chapter 1 of [52]). If for $\hat{X}, \hat{Y} \in \mathcal{N}$, there exists $\hat{Z} \in \mathcal{N}_{\hat{\rho}}$ s.t. $\hat{X} = \hat{Y} + \hat{Z}$, then we denote $\hat{X} = \hat{Y}$, $\mathbb{P}_{\hat{\rho}}$–a.s. or $\hat{\rho}$–a.s. for short.

The outcomes of the measurement of the quantum physical quantity are probabilistic in general. Let $(\Omega, \mathcal{F})$ be a certain measurable space, which describes the probabilistic events, and the map $\hat{E} : \mathcal{F} \to \mathcal{L}(\mathcal{H})$ is given. $\hat{E}$ is called a *positive operator–valued map*, which represents the instrument, if it satisfies following three conditions.

1. $\hat{E}(\Omega) = \hat{1}$
2. $\hat{E}(A) \geq 0, \forall A \in \mathcal{F}$.
3. $\hat{E}(\cup_i A_i) = \sum_i \hat{E}(A_i)$, if $A_i \in \mathcal{F}$ and $A_i \cap A_j = \emptyset$, $i \neq j$.

When one measures a quantum system and obtains stochastic event $A \in \mathcal{F}$, the density operator is updated to

$$\hat{\rho}' = \frac{\hat{M}(A)\hat{\rho}\hat{M}(A)^*}{\mathrm{Tr}[\hat{\rho}\hat{E}(A)]},$$

where $\hat{M}(A)$ is a operator–valued map satisfying $\hat{E}(A) = \hat{M}(A)^*\hat{M}(A)$. $\hat{M}$ is called a measurement operator and the updated density operator $\hat{\rho}'$ is conditional density operator on the event A. If we have a positive operator valued map $\hat{E}$, then we can define the classical probability space $(\Omega, \mathcal{F}, \mathrm{P})$ with a probability measure $\mathrm{P}(A) = \mathbb{P}_{\hat{\rho}}[\hat{E}(A)]$, for $A \in \mathcal{F}$. When we consider some quantum systems, algebraic tensor product Hilbert space is used for representation of the compound quantum system. We omit this mathematical definition (see, e.g., [12]).

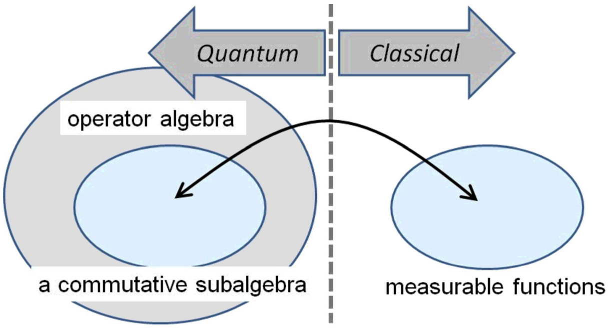

Figure 2. Quantum–classical correspondence

§ 2.2. Indirect measurement and Bayesian approach

It is difficult to observe a quantum physical quantity directly due to the measurement back actions. In this paper, we consider quantum indirect measurement to estimate the physical quantity $\hat{X}$. The quantum indirect measurement is often used in experiments and a natural setup for the quantum indirect measurement is as follows. First, we prepare a probe system $\mathcal{H}_P$ and make it interact with the system $\mathcal{H}_S$. The interaction is represented by a suitable unitary operator $\hat{U}$ over the compound system $\mathcal{H}_S \otimes \mathcal{H}_P$. For simplicity, we consider finite dimensional Hilbert spaces, i.e., $\mathcal{H}_S = \mathbb{C}^n$ and $\mathcal{H}_P = \mathbb{C}^m$. We measure a physical quantity of the probe system $\hat{Y} \in \mathcal{L}(\mathcal{H}_P)$ instead of the system's physical quantity $\hat{X} \in \mathcal{L}(\mathcal{H}_S)$. If the initial state is given by $\hat{\rho}_S \otimes \hat{\rho}_P \in \mathcal{S}(\mathcal{H}_S \otimes \mathcal{H}_P)$. After the interaction, one of eigenvalues of $\hat{U}^*(\hat{1}_S \otimes \hat{Y})\hat{U}$ is detected. Suppose the eigenvalue decomposition of $\hat{Y}$ is $\hat{Y} = \sum_{i=1}^m y_i \hat{P}(i)$, $y_i \neq y_j$ if $i \neq j$, and y_k is observed. The update of the entire density matrix is

$$\hat{\rho}_S \otimes \hat{\rho}_P \mapsto \hat{\rho}_{ent} := \frac{\hat{U}^*(\hat{1}_S \otimes \hat{P}(k))\hat{U}(\hat{\rho}_S \otimes \hat{\rho}_P)\hat{U}^*(\hat{1}_S \otimes \hat{P}(k))\hat{U}}{\mathrm{Tr}[\hat{\rho}_S \otimes \hat{\rho}_P \hat{U}^*(\hat{1}_S \otimes \hat{P}(k))\hat{U}]},$$

and the conditional expectation of the physical quantity $\hat{X}$ after the interaction is

$$\mathrm{Tr}\left[\hat{U}^*(\hat{X} \otimes \hat{1}_P)\hat{U}\hat{\rho}_{ent}\right] = \mathrm{Tr}\left[(\hat{X} \otimes \hat{1}_P)\hat{U}\hat{\rho}_{ent}\hat{U}^*\right] = \mathrm{Tr}[\hat{X}\hat{\rho}'_S],$$

where the posteriori density $\hat{\rho}'_S$ is obtained by $\hat{\rho}'_S = \mathrm{Tr}_P[\hat{U}\hat{\rho}_{ent}\hat{U}^*]$ and Tr_P is the partial trace operator that trace out over $\mathcal{H}_P$. The value $\mathrm{Tr}[\hat{X}\hat{\rho}'_S]$ is the expectation under the conditional density matrix $\hat{\rho}'_S$, so this is the quantum conditional expectation of $\hat{X}$. This Bayesian approach shows the evolution of the conditional density matrix $\hat{\rho}_S$. The update of the conditional density operator for finite dimensional descrete time quantum systems is clear, though, it is difficult to extend to the infinite dimensional continuous time quantum systems. We give another approach to the estimation of the quantum physical quantity below.

§ 2.3. A quantum minimum mean square estimation

Let $\hat{X} = \hat{X}^* \in \mathbb{C}^{n\times n}$ and a map $\hat{P} : \{1, 2, \cdots, m\} \to \mathbb{C}^{n\times n}$ satisfying $\hat{P}(i) = \hat{P}(i)^* = \hat{P}(i)^2$, $\sum_{i=1}^m \hat{P}(i) = \hat{1}$ and $\hat{P}(i)\hat{P}(j) = \delta_{ij}\hat{P}(i)$, $i, j = 1, 2, \cdots, m$, where $m \leq n$, are given. The map $\hat{P}$ is called *the projection-valued measure* or *the spectral measure* in quantum probability theory or linear operator theory. Then consider the following minimum mean square optimization problem;

$$\min_{\{(q_i,p_i)\}_{i=1}^m \subset \mathbb{R}^2} \mathbb{P}_{\hat{\rho}}\left[|\hat{X} - \hat{Q}|^2\right], \quad \hat{Q} := \sum_{i=1}^m (q_i + \mathrm{i}p_i)\hat{P}(i). \tag{2.1}$$

Since

$$\begin{aligned}
\mathbb{P}_{\hat{\rho}}\left[|\hat{X} - \hat{Q}|^2\right] =& \sum_{i\in\mathcal{I}} \mathbb{P}_{\hat{\rho}}\left[\hat{P}(i)\right]\left(q_i - \frac{1}{2}\frac{\mathbb{P}_{\hat{\rho}}\left[\hat{P}(i)\hat{X} + \hat{X}\hat{P}(i)\right]}{\mathbb{P}_{\hat{\rho}}\left[\hat{P}(i)\right]}\right)^2 \\
&+ \sum_{i\in\mathcal{I}} \mathbb{P}_{\hat{\rho}}\left[\hat{P}(i)\right]\left(p_i - \frac{1}{2\mathrm{i}}\frac{\mathbb{P}_{\hat{\rho}}\left[\hat{P}(i)\hat{X} - \hat{X}\hat{P}(i)\right]}{\mathbb{P}_{\hat{\rho}}\left[\hat{P}(i)\right]}\right)^2 \\
&+ \mathbb{P}_{\hat{\rho}}\left[\hat{X}^2\right] - \sum_{i\in\mathcal{I}} \frac{1}{4}\frac{\mathbb{P}_{\hat{\rho}}\left[\hat{P}(i)\hat{X} + \hat{X}\hat{P}(i)\right]^2}{\mathbb{P}_{\hat{\rho}}\left[\hat{P}(i)\right]} \\
&- \sum_{i\in\mathcal{I}} \frac{1}{4}\frac{\left|\mathbb{P}_{\hat{\rho}}\left[\hat{P}(i)\hat{X} - \hat{X}\hat{P}(i)\right]\right|^2}{\mathbb{P}_{\hat{\rho}}\left[\hat{P}(i)\right]},
\end{aligned}$$

where $\mathcal{I} := \{i = 1, \cdots m \mid \mathrm{Tr}[\hat{\rho}\hat{P}(i)] \neq 0\}$, the optimization problem (2.1) is easy to solve and we obtain the optimal parameters

$$q_i^{\mathrm{opt}} = \frac{1}{2}\frac{\mathbb{P}_{\hat{\rho}}\left[\hat{P}(i)\hat{X} + \hat{X}\hat{P}(i)\right]}{\mathbb{P}_{\hat{\rho}}\left[\hat{P}(i)\right]} = \mathrm{Tr}\left[\frac{\hat{\rho}\hat{P}(i) + \hat{P}(i)\hat{\rho}}{2\mathrm{Tr}[\hat{\rho}\hat{P}(i)]}\hat{X}\right], \quad \forall i \in \mathcal{I}, \tag{2.2}$$

$$p_i^{\mathrm{opt}} = \frac{1}{2\mathrm{i}}\frac{\mathbb{P}_{\hat{\rho}}\left[\hat{P}(i)\hat{X} - \hat{X}\hat{P}(i)\right]}{\mathbb{P}_{\hat{\rho}}\left[\hat{P}(i)\right]} = -\mathrm{i}\mathrm{Tr}\left[\frac{\hat{\rho}\hat{P}(i) - \hat{P}(i)\hat{\rho}}{2\mathrm{Tr}[\hat{\rho}\hat{P}(i)]}\hat{X}\right], \quad \forall i \in \mathcal{I}. \tag{2.3}$$

We define the best real and imaginary part of the optimal approximated operator as

$$\hat{Q}^{+} := \sum_{i\in\mathcal{I}} q_i^{\mathrm{opt}} \hat{P}(i), \quad \hat{Q}^{-} := \mathrm{i}\sum_{i\in\mathcal{I}} p_i^{\mathrm{opt}} \hat{P}(i), \tag{2.4}$$

and the optimal approximation is $\hat{Q}^{\mathrm{opt}} := \hat{Q}^{+} + \hat{Q}^{-}$. It is easy to confirm that the optimal approximate $\hat{Q}^{\pm}$ are characterized as follows;

$$\mathbb{P}_{\hat{\rho}}\left[\hat{Z}\left(\hat{X}-\hat{Q}^{+}\right)+\left(\hat{X}-\hat{Q}^{+}\right)\hat{Z}\right]=0, \tag{2.5}$$

$$\mathbb{P}_{\hat{\rho}}\left[\hat{Z}\hat{X}-\hat{X}\hat{Z}\right]=2\mathbb{P}_{\hat{\rho}}\left[\hat{Q}^{-}\hat{Z}\right], \quad \forall \hat{Z}\in \mathrm{alg}\left(\{\hat{P}(i)\}_{i=1}^{m}\right), \tag{2.6}$$

where $\mathrm{alg}\left(\{\hat{P}(i)\}_{j=1}^{m}\right)$ is an algebra generated by $\{\hat{P}(i)\}_{j=1}^{m}$. We define $\mathcal{Y} = \mathrm{alg}\left(\{\hat{P}_j\}_{j=1}^{m}\right)$ for short and denote $\hat{Q}^{\pm}$ by $\mathbb{Q}_{\hat{\rho}}^{\pm}[\hat{X} \mid \mathcal{Y}]$ and $\hat{Q}^{\mathrm{opt}}$ by $\mathbb{Q}_{\hat{\rho}}[\hat{X} \mid \mathcal{Y}]$. When $\hat{\rho} = |\phi\rangle\langle\phi|$ and $\hat{P}(i) = |\psi_i\rangle\langle\psi_i|$, $i = 1, 2, \cdots, n$, every optimal estimated value $q_i^{\mathrm{opt}} + \mathrm{i}p_i^{\mathrm{opt}}$ is

$$\frac{\mathrm{Tr}[\hat{\rho}\hat{P}(i)\hat{X}]}{\mathrm{Tr}[\hat{\rho}\hat{P}(i)]} = \frac{\langle\psi_i, \hat{X}\phi\rangle}{\langle\psi_i, \phi\rangle}, \quad i \in \mathcal{I},$$

where these are referred as "*weak values*" in physics [1, 18, 19, 22, 24]. If any $\hat{Z} \in \mathcal{Y}$ commute with $\hat{X}$, above condition (2.5) is the orthogonal condition under $\hat{\rho}$ and it is essentially equivalent to the classical conditional expectation. Furthermore, we can consider three interesting cases depend on the conmmutativity:

1. If $\hat{P}(i)\hat{X} = \hat{X}\hat{P}(i)$, $i \in \mathcal{I}$, or $\hat{P}(i)\hat{\rho} = \hat{\rho}\hat{P}(i)$, $i \in \mathcal{I}$, then $p_i^{\mathrm{opt}} = 0$ and

   $$q_i^{\mathrm{opt}} = \mathrm{Tr}\left[\hat{\rho}_i'\hat{X}\right], \quad \hat{\rho}_i' := \frac{\hat{P}(i)\hat{\rho}\hat{P}(i)}{\mathrm{Tr}[\hat{\rho}\hat{P}(i)]}, \quad \forall i \in \mathcal{I}.$$

   This is the result of quantum measurement. As we mention below, the filtering theory requires the commutation condition $[\hat{X}, \hat{P}(i)]_- = 0$, $i = 1, 2, \cdots, m$ [13].

2. If $\hat{X}\hat{\rho} = \hat{\rho}\hat{X}$, then $p_i^{\mathrm{opt}} = 0$ and

   $$q_i^{\mathrm{opt}} = \mathrm{Tr}\left[\hat{\rho}_i''\hat{X}\right], \quad \hat{\rho}_i'' := \frac{\sqrt{\hat{\rho}}\hat{P}(i)\sqrt{\hat{\rho}}}{\mathrm{Tr}[\hat{\rho}\hat{P}(i)]}, \quad \forall i \in \mathcal{I},$$

   where $\sqrt{\hat{\rho}}$ is the square root matrix of positive semi-definite matrix $\hat{\rho}$. The condition $\hat{X}\hat{\rho} = \hat{\rho}\hat{X}$ is equal $\mathbb{P}_{\hat{\rho}}\left[[\hat{X}, \hat{Z}]_-\right] = 0$, $\forall \hat{Z} \in \mathbb{C}^{n\times n}$. The Gammelmark's smoothing method [25] requires this condition.

3. More weakly, we can consider a condition $\mathbb{P}_{\hat{\rho}}\left[[\hat{X}, \hat{Z}]_-\right] = 0$, $\forall \hat{Z} \in \mathcal{Y}$. This condition also implies $p_i^{\mathrm{opt}} = 0$, though, this condition does not give conditional density matrix in general. A counter example is shown in Example 2.1 below.

Two obtained matrices $\hat{\rho}'_i$ and $\hat{\rho}''_i$ are also density operators, so the optimal approximation is given by "a conditional expectation." On the other hand, the general approximation (2.2) is not able to be interpreted as the conditional expectation because $\mathbb{Q}^+_{\hat{\rho}}[\hat{X}|\mathcal{Y}]$ is not positive even if $\hat{X}$ is a positive operator. For any matrix $\hat{X} \in \mathbb{C}^{n\times n}$, the following equality holds:

$$\mathbb{P}_{\hat{\rho}}\left[\hat{Z}\left(\hat{X}-\mathbb{Q}^+_{\hat{\rho}}\left[\hat{X}\mid\mathcal{Y}\right]\right)\right]=\frac{1}{2}\mathbb{P}_{\hat{\rho}}\left[\hat{Z}\hat{X}-\hat{X}\hat{Z}\right]=\mathbb{P}_{\hat{\rho}}\left[\mathbb{Q}^-_{\hat{\rho}}\left[\hat{X}\mid\mathcal{Y}\right]\hat{Z}\right],\quad\forall\hat{Z}\in\mathcal{Y}.$$

Example 2.1. Consider the following density matrix, Hermitian matrix and a set of diagonal matrices

$$\hat{\rho}=\frac{1}{2}\begin{bmatrix}1&1\\1&1\end{bmatrix},\quad\hat{X}=\begin{bmatrix}1&2-\mathrm{i}\alpha\\2+\mathrm{i}\alpha&3\end{bmatrix},\ \alpha\in\mathbb{R},\ \mathcal{Y}=\left\{\begin{bmatrix}a&0\\0&b\end{bmatrix},\ a,b\in\mathbb{C}\right\}.$$

Then,

$$\mathbb{Q}^+_{\hat{\rho}}\left[\hat{X}|\mathcal{Y}\right]=\begin{bmatrix}3&0\\0&5\end{bmatrix},\quad\mathbb{Q}^-_{\hat{\rho}}\left[\hat{X}|\mathcal{Y}\right]=\mathrm{i}\alpha\begin{bmatrix}-1&0\\0&1\end{bmatrix}.$$

The matrix $\hat{X}$ has a negative eigenvalue, though, the approximation $\mathbb{Q}^+_{\hat{\rho}}[\hat{X}|\mathcal{Y}]$ is a positive–definite matrix. If $\alpha=0$, then $\mathbb{P}_{\hat{\rho}}\left[\left[\hat{X},\hat{Z}\right]_-\right]=0$, $\forall\hat{Z}\in\mathcal{Y}$.

If we choose the $\hat{X}'$ as a positive–definite matrix

$$\hat{X}'=\begin{bmatrix}\sqrt{1.11}-0.1&-1\\-1&\sqrt{1.11}+0.1\end{bmatrix},$$

then

$$\mathbb{Q}^+_{\hat{\rho}}\left[\hat{X}'|\mathcal{Y}\right]=\frac{1}{2}\begin{bmatrix}\sqrt{1.11}-1.1&0\\0&\sqrt{1.11}+1.1\end{bmatrix}.$$

This approximation has negative eigenvalue, so positive–definiteness is not preserved in general.

§ 3. Quantum conditional expectation and minimum mean square approximation

In this section, we introduce real and imaginary parts of the best approximation in the sense of the semi-norms induced by the pre-inner products below and show several properties of them. Let $\mathcal{Y}$ be a commutative $*$–subalgebra of $\mathcal{L}(\mathcal{H})$. We introduce another $*$–subalgebra whose elements commute with all of the elements in $\mathcal{Y}$;

$$\mathcal{Y}':=\{\hat{X}\in\mathcal{L}(\mathcal{H})\mid\hat{X}\hat{Y}=\hat{Y}\hat{X},\ \forall\hat{Y}\in\mathcal{Y}\}.$$

Hereafter we assume $\mathcal{Y} = (\mathcal{Y}')'$, i.e., $\mathcal{Y}$ is a commutative von Neumann subalgebra [13, 52]. For instance, $\mathcal{Y} = \mathrm{alg}(\{\hat{P}_j\}_{j=1}^m)$ is a commutative von Neumann subalgebra of $\mathbb{C}^{n\times n}$. von Neumann algebras are a generalization of the set of the σ–measurable bounded functions and especially a commutative von Neumann algebra is isomorphic to the set of the σ–measurable bounded functions. Note that $\mathcal{Y}'$ is generally non-commutative $*$–subalgebra. Let us define the quantum conditional expectation (see, e.g., [52, Prop. 2.36] and [13, Sec. 3]) and the optimal approximation as we discussed above.

§ 3.1. Definitions

We introduce three approximations of a given $\hat{X} \in \mathcal{L}(\mathcal{H})$. All of them are based on the following pre-inner products [4].

Definition 3.1. For given $\hat{\rho} \in \mathcal{S}(\mathcal{H})$,

1. the pre–inner product $\langle \bullet, \bullet \rangle_{\hat{\rho}} : \mathcal{L}(\mathcal{H}) \times \mathcal{L}(\mathcal{H}) \to \mathbb{C}$ is defined by $\langle \hat{X}, \hat{Y} \rangle_{\hat{\rho}} := \mathbb{P}_{\hat{\rho}}\left[\hat{X}^*\hat{Y}\right]$.
2. the symmetric pre–inner product $\langle\langle \bullet, \bullet \rangle\rangle_{\hat{\rho}} : \mathcal{L}(\mathcal{H}) \times \mathcal{L}(\mathcal{H}) \to \mathbb{C}$ is defined by $\langle\langle \hat{X}, \hat{Y} \rangle\rangle_{\hat{\rho}} := \frac{1}{2}\mathbb{P}_{\hat{\rho}}\left[\hat{X}^*\hat{Y} + \hat{Y}\hat{X}^*\right]$.

Note that $\langle\langle \hat{X}, \hat{Y} \rangle\rangle_{\hat{\rho}}$ is not a real part of $\langle \hat{X}, \hat{Y} \rangle_{\hat{\rho}}$. The pre-inner product $\langle\langle \bullet, \bullet \rangle\rangle_{\hat{\rho}}$ is also used in quantum infromation geometry [4]. These pre–inner products satisfy the Cauchy–Schwarz inequality (see, e.g., Proposition 9.5 of [52]). $\langle \hat{X}, \hat{X} \rangle_{\hat{\rho}} = 0$ is necessary and sufficient condition for $\hat{X} \in \mathcal{N}_{\hat{\rho}}$, though, $\langle\langle \hat{X}, \hat{X} \rangle\rangle_{\hat{\rho}} = 0$ is not. If $\hat{X} \in \mathcal{N}_{\hat{\rho}} \cap \mathcal{Y}$, then $\langle \hat{X}, \hat{X} \rangle_{\hat{\rho}} = \langle\langle \hat{X}, \hat{X} \rangle\rangle_{\hat{\rho}} = 0$. This is proven by the Cauchy–Schwarz inequality and commutativity of $\mathcal{Y}$. We use two measures to find the best approximation in $\mathcal{Y}$, where the "probability zero" space is common whenever any of two semi–inner products is used.

Definition 3.2 (Quantum conditional expectation).

Let $(\mathcal{L}(\mathcal{H}), \mathbb{P}_{\hat{\rho}})$ be a quantum probability space and $\mathcal{Y}$ be a commutative von Neumann sub–algebra of $\mathcal{L}(\mathcal{H})$. A linear operator $\hat{Q} \in \mathcal{Y}$ is called a version of the quantum conditional expectation if there exists $\hat{Q} \in \mathcal{Y}$ satisfies

$$\langle \hat{Z}, \hat{X} - \hat{Q} \rangle_{\hat{\rho}} = 0, \quad \forall \hat{Z} \in \mathcal{Y} \tag{3.1}$$

for arbitrary fixed $\hat{X} \in \mathcal{Y}'$. Then we denote $\hat{Q} = \mathbb{P}_{\hat{\rho}}\left[\hat{X}|\mathcal{Y}\right]$.

Some properties of the quantum conditional expectation are shown in, for example, [13]. The definition of the quantum conditional expectation implies that the $\hat{X} - \mathbb{P}_{\hat{\rho}}\left[\hat{X}|\mathcal{Y}\right]$ and the commutative sub-algebra $\mathcal{Y}$ are orthogonal under the state $\mathbb{P}_{\hat{\rho}}$. We extend the definition of orthogonality to non-commutative regime.

Definition 3.3.

Let $(\mathcal{L}(\mathcal{H}), \mathbb{P}_{\hat{\rho}})$ be a quantum probability space and $\mathcal{Y}$ be a commutative von Neumann sub–algebra of $\mathcal{L}(\mathcal{H})$. For arbitrary fixed $\hat{X} \in \mathcal{L}(\mathcal{H})$, we define following operators:

1. A linear operator $\hat{Q} \in \mathcal{Y}$ is called a version of symmetric quantum least mean square approximation if there exists $\hat{Q} \in \mathcal{Y}$ that satisfies

   $$\langle\langle \hat{Z}, \hat{X} - \hat{Q} \rangle\rangle_{\hat{\rho}} = 0, \quad \forall \hat{Z} \in \mathcal{Y}. \tag{3.2}$$

   Then we denote $\mathbb{Q}^{+}_{\hat{\rho}} \left[\hat{X}|\mathcal{Y}\right] = \hat{Q}$.

2. A linear operator $\hat{Q} \in \mathcal{Y}$ is called a version of the mean non-commutativity with respect to $\mathcal{Y}$ if there exists $\hat{Q} \in \mathcal{Y}$ that satisfies

   $$\mathbb{P}_{\hat{\rho}} \left[\hat{Z}\hat{X} - \hat{X}\hat{Z}\right] = 2\mathbb{P}_{\hat{\rho}} \left[\hat{Q}\hat{Z}\right], \quad \forall \hat{Z} \in \mathcal{Y}. \tag{3.3}$$

   Then we denote $\mathbb{Q}^{-}_{\hat{\rho}} \left[\hat{X}|\mathcal{Y}\right] = \hat{Q}$.

3. $\mathbb{Q}_{\hat{\rho}} \left[\hat{X}|\mathcal{Y}\right] := \mathbb{Q}^{+}_{\hat{\rho}} \left[\hat{X}|\mathcal{Y}\right] + \mathbb{Q}^{-}_{\hat{\rho}} \left[\hat{X}|\mathcal{Y}\right]$ is called the minimum mean square estimation of $\hat{X}$ with respect to $\mathcal{Y}$.

If $\mathcal{N}_{\hat{\rho}} \cap \mathcal{Y} \neq \{0\}$, then there are many operators that satisfy above conditions. This is why we use "a version of" here. We call Eq. (3.2) *the symmetric orthogonal condition*. This is not a quantum conditional expectation in the sense of Takesaki's requirements for quantum conditional expectation [52]. Obviously, $\mathbb{P}_{\hat{\rho}} \left[\hat{X}\right] = \mathbb{P}_{\hat{\rho}} \left[\mathbb{Q}_{\hat{\rho}} \left[\hat{X}|\mathcal{Y}\right]\right] = \mathbb{P}_{\hat{\rho}} \left[\mathbb{Q}^{+}_{\hat{\rho}} \left[\hat{X}|\mathcal{Y}\right]\right]$ holds, i.e., these two approximations are unbiased estimates. The name "the symmetric minimum square approximation" is originated from Proposition 3.5.

Since the expectation of $\mathbb{P}_{\hat{\rho}} \left[\mathbb{Q}^{-}_{\hat{\rho}} \left[\hat{X}|\mathcal{Y}\right]\right]$ is always zero, it is difficult to find its statistical meaning. However, this is an interesting quantity in the view of non-commutative geometry. If $\hat{X}$ and $\hat{Z}$ are Hilbert–Schmidt class operators, respectively, $[\hat{X}, \hat{Z}]_{-}$ is orthogonal to both of $\hat{X}$ and $\hat{Z}$ in the sense of Hilbert–Schmidt inner product $\langle \bullet, \bullet \rangle_{\hat{\rho}}$. From Eq. (2.3), the operator $\mathbb{Q}^{-}_{\hat{\rho}}[\hat{X}|\mathcal{Y}]$ is a measure of the "$\hat{\rho}$–direction" component of the orthogonal direction against to the both of $\hat{X}$ and $\mathcal{Y}$.

§ 3.2. Basic properties

A list of the basic properties of $\mathbb{Q}^{\pm}_{\hat{\rho}} \left[\hat{X} \mid \mathcal{Y}\right]$ is as follows:

1. **(linearity)** $\mathbb{Q}^{\pm}_{\hat{\rho}} \left[\hat{X} \mid \mathcal{Y}\right]$ is linear in $\hat{X} \in \mathcal{L}(\mathcal{H})$.

2. **(uniqueness)** $\mathbb{Q}^{\pm}_{\hat{\rho}} \left[\hat{X} \mid \mathcal{Y}\right]$ is uniquely determined in the sense of $\mathbb{P}_{\hat{\rho}}$–a.s.

3. **(self-adjointness and skewness)** $\mathbb{Q}_{\hat\rho}^{\pm}\left[\hat X \mid \mathcal{Y}\right]^{*} = \pm\mathbb{Q}_{\hat\rho}^{\pm}\left[\hat X \mid \mathcal{Y}\right]$, $\mathbb{P}_{\hat\rho}$–a.s. for $\hat X = \hat X^{*} \in \mathcal{L}(\mathcal{H})$, and $\mathbb{Q}_{\hat\rho}^{\pm}\left[\hat X \mid \mathcal{Y}\right]^{*} = \mp\mathbb{Q}_{\hat\rho}^{\pm}\left[\hat X \mid \mathcal{Y}\right]$, $\mathbb{P}_{\hat\rho}$–a.s. for $\hat X = -\hat X^{*} \in \mathcal{L}(\mathcal{H})$.

The proofs of above properties are trivial from their definition, and we omit the proofs. From above properties, $\mathbb{Q}_{\hat\rho}^{+}\left[\hat X \mid \mathcal{Y}\right]$ is the minimum mean square error approximation for $\hat X = \hat X^{*}$ in self-adjoint operators in $\mathcal{Y}$ as follows.

Proposition 3.4 (MMSE approximation in $\langle\langle\bullet,\bullet\rangle\rangle$ sense [43]).
For arbitrary $\hat X \in \mathcal{L}(\mathcal{H})$,

$$\langle\langle \hat X - \mathbb{Q}_{\hat\rho}^{+}[\hat X|\mathcal{Y}], \hat X - \mathbb{Q}_{\hat\rho}^{+}[\hat X|\mathcal{Y}]\rangle\rangle_{\hat\rho} \leq \langle\langle \hat X - \hat Z, \hat X - \hat Z\rangle\rangle_{\hat\rho}, \quad \forall \hat Z \in \mathcal{Y}.$$

Since $\langle \hat X - \hat Z, \hat X - \hat Z\rangle_{\hat\rho} = \langle\langle \hat X - \hat Z, \hat X - \hat Z\rangle\rangle_{\hat\rho}$ for $\hat X = \hat X^{*}$ and $\hat Z = \hat Z^{*}$, the following inequality also holds.

$$\langle \hat X - \mathbb{Q}_{\hat\rho}^{+}[\hat X|\mathcal{Y}], \hat X - \mathbb{Q}_{\hat\rho}^{+}[\hat X|\mathcal{Y}]\rangle_{\hat\rho} \leq \langle \hat X - \hat Z, \hat X - \hat Z\rangle_{\hat\rho}, \quad \forall \hat Z = \hat Z^{*} \in \mathcal{Y}.$$

According to the previous section, $\mathbb{Q}_{\hat\rho}\left[\hat X|\mathcal{Y}\right] := \mathbb{Q}_{\hat\rho}^{+}\left[\hat X|\mathcal{Y}\right] + \mathbb{Q}_{\hat\rho}^{-}\left[\hat X|\mathcal{Y}\right]$ is the best approximation in finite dimensional case. It is also true if we consider bounded operators on general separable Hilbert space.

Proposition 3.5 (MMSE approximation in $\langle\bullet,\bullet\rangle$ sense [43]).
For arbitrary $\hat X = \hat X^{*} \in \mathcal{L}(\mathcal{H})$,

$$\langle \hat X - \mathbb{Q}_{\hat\rho}[\hat X|\mathcal{Y}], \hat X - \mathbb{Q}_{\hat\rho}[\hat X|\mathcal{Y}]\rangle_{\hat\rho} \leq \langle \hat X - \hat Z, \hat X - \hat Z\rangle_{\hat\rho}, \quad \forall \hat Z \in \mathcal{Y}.$$

Furthermore, the approximation error is

$$\langle \hat X - \mathbb{Q}_{\hat\rho}[\hat X|\mathcal{Y}], \hat X - \mathbb{Q}_{\hat\rho}[\hat X|\mathcal{Y}]\rangle_{\hat\rho} = \langle \hat X, \hat X\rangle_{\hat\rho} - \langle \mathbb{Q}_{\hat\rho}[\hat X|\mathcal{Y}], \mathbb{Q}_{\hat\rho}[\hat X|\mathcal{Y}]\rangle_{\hat\rho}. \tag{3.4}$$

One of our interest is whether richer information gives better estimation or not in quantum estimation theory. Proposition 3.5 gives the following results.

Corollary 3.6 ([43]).
Let $\mathcal{Y}_1$ and $\mathcal{Y}_2$ be commutative $*$–subalgebras of $\mathcal{L}(\mathcal{H})$ and have inclusion relation $\mathcal{Y}_1 \subseteq \mathcal{Y}_2$. Then, for any $\hat X = \hat X^{*} \in \mathcal{L}(\mathcal{H})$,

1. $\langle \hat X - \mathbb{Q}_{\hat\rho}^{+}[\hat X|\mathcal{Y}_2], \hat X - \mathbb{Q}_{\hat\rho}^{+}[\hat X|\mathcal{Y}_2]\rangle_{\hat\rho} \leq \langle \hat X - \mathbb{Q}_{\hat\rho}^{+}[\hat X|\mathcal{Y}_1], \hat X - \mathbb{Q}_{\hat\rho}^{+}[\hat X|\mathcal{Y}_1]\rangle_{\hat\rho}$.
2. $\langle \hat X - \mathbb{Q}_{\hat\rho}^{-}[\hat X|\mathcal{Y}_2], \hat X - \mathbb{Q}_{\hat\rho}^{-}[\hat X|\mathcal{Y}_2]\rangle_{\hat\rho} \leq \langle \hat X - \mathbb{Q}_{\hat\rho}^{-}[\hat X|\mathcal{Y}_1], \hat X - \mathbb{Q}_{\hat\rho}^{-}[\hat X|\mathcal{Y}_1]\rangle_{\hat\rho}$.

In similar way, we obtain the following corollary.

Corollary 3.7. *Let $\mathcal{Y}_1$ and $\mathcal{Y}_2$ be commutative $*$–subalgebras of $\mathcal{L}(\mathcal{H})$ and have inclusion relation $\mathcal{Y}_1 \subseteq \mathcal{Y}_2$. Then, for any $\hat{X} = \hat{X}^* \in \mathcal{L}(\mathcal{H})$,*

$$\langle \hat{X} - \mathbb{Q}_{\hat{\rho}}[\hat{X}|\mathcal{Y}_2], \hat{X} - \mathbb{Q}_{\hat{\rho}}[\hat{X}|\mathcal{Y}_2]\rangle_{\hat{\rho}} \leq \langle \hat{X} - \mathbb{Q}_{\hat{\rho}}[\hat{X}|\mathcal{Y}_1], \hat{X} - \mathbb{Q}_{\hat{\rho}}[\hat{X}|\mathcal{Y}_1]\rangle_{\hat{\rho}}.$$

§ 3.3. Some lower bounds and remarks

To estimate the approximation error bound is important for accuracy. Roles of the $\mathbb{Q}^-_{\hat{\rho}}[\hat{X}|\mathcal{Y}]$ is still unclear for the author, though, it provides a lower bound of real MMSE and is used in the quantum smoothing equation introduced below. The following proposition holds.

Proposition 3.8 (A lower bound of MMSE).
For $\hat{X} = \hat{X}^* \in \mathcal{L}(\mathcal{H})$,

$$\mathbb{P}_{\hat{\rho}}\left[\left(\hat{X} - \mathbb{Q}^+_{\hat{\rho}}[\hat{X}|\mathcal{Y}]\right)^2\right] \geq \mathbb{P}_{\hat{\rho}}\left[\left|\mathbb{Q}^-_{\hat{\rho}}[\hat{X}|\mathcal{Y}]\right|^2\right], \tag{3.5}$$

where $|\hat{A}|^2 := \hat{A}^*\hat{A}$.

Proof. From the definitions,

$$\langle \hat{Z}, \hat{X} - \mathbb{Q}^+_{\hat{\rho}}[\hat{X}|\mathcal{Y}]\rangle_{\hat{\rho}} = \langle \hat{Z}, \mathbb{Q}^-_{\hat{\rho}}[\hat{X}|\mathcal{Y}]\rangle_{\hat{\rho}}, \quad \forall \hat{Z} \in \mathcal{Y}$$

holds. Since $\mathbb{Q}^-_{\hat{\rho}}[\hat{X}|\mathcal{Y}] \in \mathcal{Y}$, we choose $\hat{Z} = \mathbb{Q}^-_{\hat{\rho}}[\hat{X}|\mathcal{Y}]$ and then

$$\langle \mathbb{Q}^-_{\hat{\rho}}[\hat{X}|\mathcal{Y}], \hat{X} - \mathbb{Q}^+_{\hat{\rho}}[\hat{X}|\mathcal{Y}]\rangle_{\hat{\rho}} = \langle \mathbb{Q}^-_{\hat{\rho}}[\hat{X}|\mathcal{Y}], \mathbb{Q}^-_{\hat{\rho}}[\hat{X}|\mathcal{Y}]\rangle_{\hat{\rho}} \geq 0.$$

Using Schwarz's inequality, we obtain

$$\langle \mathbb{Q}^-_{\hat{\rho}}[\hat{X}|\mathcal{Y}], \hat{X} - \mathbb{Q}^+_{\hat{\rho}}[\hat{X}|\mathcal{Y}]\rangle_{\hat{\rho}} \leq \sqrt{\mathbb{P}_{\hat{\rho}}\left[\left|\mathbb{Q}^-_{\hat{\rho}}[\hat{X}|\mathcal{Y}]\right|^2\right]\mathbb{P}_{\hat{\rho}}\left[\left(\hat{X} - \mathbb{Q}^+_{\hat{\rho}}[\hat{X}|\mathcal{Y}]\right)^2\right]}$$

and the inequality (3.5) holds. □

For example, consider the MMSE of Example 2.1. In this case, the following equality holds;

$$\mathbb{P}_{\hat{\rho}}\left[\left(\hat{X} - \mathbb{Q}^+_{\hat{\rho}}[\hat{X}|\mathcal{Y}]\right)^2\right] = \alpha^2 = \mathbb{P}_{\hat{\rho}}\left[\left|\mathbb{Q}^-_{\hat{\rho}}[\hat{X}|\mathcal{Y}]\right|^2\right].$$

Here, $\mathbb{Q}^-_{\hat{\rho}}[\hat{X}|\mathcal{Y}]$ is a measure of non-commutativity in estimation and α is magnitude of non-commutativity between $\hat{X}$ and $\mathcal{Y}$ under the given state $\mathbb{P}_{\hat{\rho}}$.

The lower bound (3.5) is not tight in general. If $\hat{X} \in \mathcal{Y}'$, then $\mathbb{Q}^-_{\hat{\rho}}[\hat{X}|\mathcal{Y}] = 0$ but this does not imply that the MMSE is sufficiently small. It is known well that self-adjoint operators on the two dimensional Hilbert space can be represented by three dimensional real space [41]. We, then, can illustrate the approximation of an operator $\hat{X}$ as in Fig. 3. We represent the vertical axis is for the self adjoint operators in $\mathcal{Y}$ and the rotation of the vertical axis is for skew operators in $\mathcal{Y}$. The amplitude of the rotation $\mathbb{Q}^-_{\hat{\rho}}[\hat{X}|\mathcal{Y}]$ is the Euclid distance between the vertical axis and $\hat{X}$.

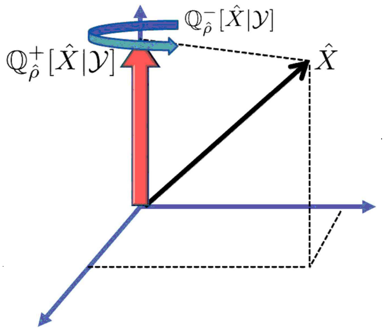

Figure 3. A interpretation of the estimation

In quantum theory, the famous estimation accuracy called uncertainty relation is well known. A kind of uncertainty relation holds for the MMSE approximates.

Proposition 3.9 (Uncertainty relation).
For any $\hat{X}_1 = \hat{X}_1^*$, $\hat{X}_2 = \hat{X}_2^* \in \mathcal{L}(\mathcal{H})$, *define* $\Delta\hat{X}_i := \hat{X}_i - \mathbb{Q}^+_{\hat{\rho}}[\hat{X}_i \mid \mathcal{Y}]$, $i = 1, 2$.

(3.6)

$$\mathbb{P}_{\hat{\rho}}\left[\left(\Delta\hat{X}_1\right)^2\right]\mathbb{P}_{\hat{\rho}}\left[\left(\Delta\hat{X}_2\right)^2\right] - \frac{1}{4}\mathbb{P}_{\hat{\rho}}\left[\left([\Delta\hat{X}_1, \Delta\hat{X}_2]_+\right)\right]^2$$
$$\geq \left|\frac{1}{2}\mathbb{P}_{\hat{\rho}}\left[[\hat{X}_1, \hat{X}_2]_-\right] + \langle\mathbb{Q}^-_{\hat{\rho}}[\hat{X}_1 \mid \mathcal{Y}]\ ,\ \mathbb{Q}^+_{\hat{\rho}}[\hat{X}_2 \mid \mathcal{Y}]\rangle_{\hat{\rho}} + \langle[\mathbb{Q}^+_{\hat{\rho}}[\hat{X}_1 \mid \mathcal{Y}]\ ,\ \mathbb{Q}^-_{\hat{\rho}}[\hat{X}_2 \mid \mathcal{Y}]\rangle_{\hat{\rho}}\right|^2.$$

Proof. Consider the covariance matrix

$$\begin{aligned}
&\mathbb{P}_{\hat{\rho}}\left[\begin{bmatrix}\Delta\hat{X}_1 \\ \Delta\hat{X}_2\end{bmatrix}\begin{bmatrix}\Delta\hat{X}_1 & \Delta\hat{X}_2\end{bmatrix}\right] \\
&= \begin{bmatrix}\mathbb{P}_{\hat{\rho}}[\Delta\hat{X}_1^2] & \mathbb{P}_{\hat{\rho}}[\Delta\hat{X}_1\Delta\hat{X}_2] \\ \mathbb{P}_{\hat{\rho}}[\Delta\hat{X}_2\Delta\hat{X}_1] & \mathbb{P}_{\hat{\rho}}[\Delta\hat{X}_2^2]\end{bmatrix} \\
&= \begin{bmatrix}\mathbb{P}_{\hat{\rho}}[\Delta\hat{X}_1^2] & \frac{1}{2}\mathbb{P}_{\hat{\rho}}[[\Delta\hat{X}_1, \Delta\hat{X}_2]_+] \\ \frac{1}{2}\mathbb{P}_{\hat{\rho}}[[\Delta\hat{X}_1, \Delta\hat{X}_2]_+] & \mathbb{P}_{\hat{\rho}}[\Delta\hat{X}_2^2]\end{bmatrix} + \frac{\mathbb{P}_{\hat{\rho}}[[\Delta\hat{X}_1, \Delta\hat{X}_2]_-]}{2}\begin{bmatrix}0 & 1 \\ -1 & 0\end{bmatrix} \geq 0.
\end{aligned}$$

The determinant of the semi–positive definite matrix is non-negative. The straight calculation gives

$$\mathbb{P}_{\hat{\rho}}[\Delta\hat{X}_1^2]\mathbb{P}_{\hat{\rho}}[\Delta\hat{X}_2^2] - \frac{1}{4}\mathbb{P}_{\hat{\rho}}[[\Delta\hat{X}_1, \Delta\hat{X}_2]_+]^2 + \frac{1}{4}\mathbb{P}_{\hat{\rho}}[[\Delta\hat{X}_1, \Delta\hat{X}_2]_-]^2 \geq 0$$

Since

$$\begin{aligned}\frac{1}{2}\mathbb{P}_{\hat{\rho}}[[\Delta\hat{X}_1, \Delta\hat{X}_2]_-] =& \frac{1}{2}\mathbb{P}_{\hat{\rho}}[[\hat{X}_1, \hat{X}_2]_-] - \mathbb{P}_{\hat{\rho}}[\mathbb{Q}_{\hat{\rho}}^-[\hat{X}_1 \mid \mathcal{Y}]\mathbb{Q}_{\hat{\rho}}^+[\hat{X}_2 \mid \mathcal{Y}]] \\ &+ \mathbb{P}_{\hat{\rho}}[\mathbb{Q}_{\hat{\rho}}^+[\hat{X}_1 \mid \mathcal{Y}]\mathbb{Q}_{\hat{\rho}}^-[\hat{X}_2 \mid \mathcal{Y}]] \\ =& \frac{1}{2}\mathbb{P}_{\hat{\rho}}[[\hat{X}_1, \hat{X}_2]_-] + \langle \mathbb{Q}_{\hat{\rho}}^-[\hat{X}_1 \mid \mathcal{Y}] \,,\, \mathbb{Q}_{\hat{\rho}}^+[\hat{X}_2 \mid \mathcal{Y}]\rangle_{\hat{\rho}} \\ &+ \langle [\mathbb{Q}_{\hat{\rho}}^+[\hat{X}_1 \mid \mathcal{Y}] \,,\, \mathbb{Q}_{\hat{\rho}}^-[\hat{X}_2 \mid \mathcal{Y}]]\rangle_{\hat{\rho}}\end{aligned}$$

and $\mathbb{P}_{\hat{\rho}}[[\Delta\hat{X}_1, \Delta\hat{X}_2]_-]^2 = -|\mathbb{P}_{\hat{\rho}}[[\Delta\hat{X}_1, \Delta\hat{X}_2]_-]|^2$, the inequality (3.6) holds.

□

The inequality (3.6) is a generalization of the Schrödinger–Robertson type uncertainty relation [35]. From the view of quantum estimation theory [59], the uncertainty principle gives a physical estimation accuracy bound. When $\mathbb{Q}_{\hat{\rho}}^-[\hat{X}_i|\mathcal{Y}] = 0$, $i = 1, 2$, the inequality (3.6) is same as the Schrödinger–Robertson type uncertainty relation. If we want to beat the Schrödinger–Robertson type uncertainty relation of physical quantities $\hat{X}_1$ and $\hat{X}_2$, we should choose the measurement algebra $\mathcal{Y}$ to avoid the condition $\mathbb{Q}_{\hat{\rho}}^-[\hat{X}_i|\mathcal{Y}] = 0$, $i = 1, 2$ simultaneously.

Remark.

The real minimum mean square approximate of $\hat{X} \in \mathcal{L}(\mathcal{H})$ does not satisfy the orthogonality condition (3.1). From the definition of the minimum square approximation (3.2), the "normal" orthogonal relation becomes

$$\mathbb{P}_{\hat{\rho}}\left[\hat{Z}\left(\hat{X} - \mathbb{Q}_{\hat{\rho}}^+[\hat{X}|\mathcal{Y}]\right)\right] = \mathbb{P}_{\hat{\rho}}\left[\mathbb{Q}_{\hat{\rho}}^-[\hat{X}|\mathcal{Y}]\hat{Z}\right], \ \forall\hat{Z} \in \mathcal{Y}.$$

Obviously, the estimation error $\hat{X} - \mathbb{Q}_{\hat{\rho}}^+[\hat{X}|\mathcal{Y}] \in \mathcal{L}(\mathcal{H})$ is orthogonal to $\mathcal{Y}$ if and only if $\mathbb{Q}_{\hat{\rho}}^-[\hat{X}|\mathcal{Y}] = 0$ under $\mathbb{P}_{\hat{\rho}}$. $\hat{X} \in \mathcal{Y}'$ is a sufficient condition for the orthogonality condition (3.1).

Remark.

$\mathbb{P}_{\hat{\rho}}[\bullet|\mathcal{Y}]$, and $\mathbb{Q}_{\hat{\rho}}^{\pm}[\bullet|\mathcal{Y}]$ are regarded as a linear functional on their domains because any commutative $*$–subalgebra can be seemed as measurable functions on a suitable chosen measurable space $(\Omega, \mathcal{F})$. There exists a $*$–isomorphism ι between $\mathcal{Y}$ and $L^\infty(\Omega)$.

§ 4. Model and quantum filtering

§ 4.1. Model

Any quantum system is described by suitable Hilbert space and linear operators on the Hilbert space. We consider two quantum systems, system and probe system. We describe them $\mathcal{H}_S$ and $\mathcal{H}_P$, respectively. $\mathcal{H}_P$ is a continuous Fock space [27]; $\mathcal{H}_P = \otimes_{t\in[0,\infty)}\mathcal{H}_P(t)$. The compound quantum system is the tensor product Hilbert space $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_P$ equipped with a density operator $\hat{\rho} = \hat{\rho}_S \otimes \hat{\rho}_P$, $\hat{\rho}_S \in \mathcal{S}(\mathcal{H}_S)$, $\hat{\rho}_P \in \mathcal{S}(\mathcal{H}_P)$. Physical quantities of the system are described by self-adjoint operators in $\mathcal{L}(\mathcal{H}_S)$ and physical quantities of the probe system are described by self-adjoint operators in $\mathcal{L}(\mathcal{H}_P)$. They act on the total quantum system with corresponding identity operator, though, we omit identity operator for simplicity; $\hat{X}\otimes\hat{1}_P \equiv \hat{X}$ and $\hat{1}_S\otimes\hat{Y} \equiv \hat{Y}$ for $\hat{X} \in \mathcal{L}(\mathcal{H}_S)$ and $\hat{Y} \in \mathcal{L}(\mathcal{H}_P)$.

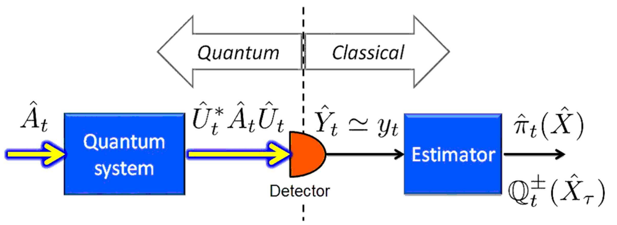

Figure 4. Schematic diagram

In order to quantum theory, the time evolution of every physical quantity $\hat{X} = \hat{X}^* \in \mathcal{L}(\mathcal{H})$ driven by probe system is determined by a unitary operator $\hat{U}_t$ that describes the interaction between the system and the probe. We consider the unitary operator $\hat{U}_t$ as the solution of the following equation;

$$\frac{d}{dt}\hat{U}_t = \left(-\mathrm{i}\hat{H} + \hat{L}\hat{a}_t^* - \hat{L}^*\hat{a}_t\right)\hat{U}_t, \quad \hat{U}_0 = \hat{1} \tag{4.1}$$

where $\hat{a}_t \in \mathcal{L}(\mathcal{H}_P(t))$ is called the quantum white noise which satisfies

$$[\hat{a}_t, \hat{a}_s^*]_- = \delta(t-s)\hat{1},$$

where δ is Dirac's delta function. The formal integral of the quantum white noise and its infinitesimal increment are defined

$$\hat{A}_t := \int_0^t \hat{a}_s ds, \quad d\hat{A}_t = \hat{A}(t+dt) - \hat{A}(t),$$

respectively [26]. In order to quantum stochastic calculus [36], the quantum Ito's rule is

$$\begin{cases} d\hat{A}_t d\hat{A}_t = d\hat{A}_t^* d\hat{A}_t = d\hat{A}_t dt = (dt)^2 = 0, \\ d\hat{A}_t d\hat{A}_t^* = dt \end{cases} . \tag{4.2}$$

From Wong–Zakai's theorem, the formal equation (4.1) is described by *the Hudson–Parthasarathy equation*

$$d\hat{U}_t = \left(-\mathrm{i}\hat{H}dt - \frac{1}{2}\hat{L}^*\hat{L}dt + \hat{L}d\hat{A}_t^* - \hat{L}^* d\hat{A}_t \right) \hat{U}_t. \tag{4.3}$$

Then the time evolution of the $\hat{X}_t = \hat{U}_t^* \hat{X} \hat{U}_t$ follows the quantum stochastic differential equation

$$\begin{aligned} d\hat{X}_t =& \mathrm{i}[\hat{H}_t, \hat{X}_t]_- dt + \frac{1}{2}\left(\hat{L}_t^*[\hat{X}_t, \hat{L}_t]_- + [\hat{L}_t^*, \hat{X}_t]_- \hat{L}_t \right) dt \\ &+ [\hat{L}_t^*, \hat{X}_t]_- d\hat{A}_t + [\hat{X}_t, \hat{L}_t]_- d\hat{A}_t^*, \end{aligned} \tag{4.4}$$

where $\hat{H}_t = \hat{U}_t^* \hat{H} \hat{U}_t$, $\hat{L}_t = \hat{U}_t^* \hat{L} \hat{U}_t$, and $[\hat{A}, \hat{B}]_- := \hat{A}\hat{B} - \hat{B}\hat{A}$. For derivation, see [13, 16, 26, 61].

We consider the balanced homodyne detection as a detection of the probe system. Its POVM is introduced in [7] and the dynamical representation is in, for example, [26, 58]. The measurement outcome is $\hat{Y}_t = \hat{U}_t^*(\hat{A}_t + \hat{A}_t^*)\hat{U}_t$ and its increment is

$$d\hat{Y}_t = \left(\hat{L}_t + \hat{L}_t^* \right) dt + d\hat{A}_t + d\hat{A}_t^*. \tag{4.5}$$

We define the following $*$–algebra by double commutant of the measurement records;

$$\mathcal{Y}_t := \left(\{\hat{Y}_s;\ 0 \le s \le t\}' \right)'.$$

From the definitions of the unitary operator and the observed process, following equations hold.

$$\hat{X}_t \hat{Y}_s = \hat{Y}_s \hat{X}_t, \quad \forall t \ge s \ge 0, \tag{4.6}$$

$$\hat{Y}_t \hat{Y}_s = \hat{Y}_s \hat{Y}_t, \quad \forall t, s \ge 0. \tag{4.7}$$

These ensure that $\mathcal{Y}_t$ is a commutative von Neumann subalgebra and $\hat{X}_t \in \mathcal{Y}_t'$ for $t \ge 0$. $\mathcal{Y}_t$ is the quantum counter part of $\sigma(y_s;\ 0 \le s \le t)$ where y_t is a classical signal.

We use following lemma in order to derive the filtering and smoothing equations; see, for example, [13].

Lemma 4.1.

1. $\mathcal{Y}_s \subseteq \mathcal{Y}_t$ *for* $t \geq s$.

2. $\mathcal{Y}'_t \subseteq \mathcal{Y}'_s$ *for* $t \geq s$.

3. ***(tower property)*** $\mathbb{P}_{\hat{\rho}}[\mathbb{P}_{\hat{\rho}}[\hat{X}|\mathcal{Y}_t]|\mathcal{Y}_s] = \mathbb{P}_{\hat{\rho}}[\hat{X}|\mathcal{Y}_s]$ *for* $t \geq s$ *and* $\hat{X} \in \mathcal{Y}'_t$.

$\hat{X}_\tau$ lies in $\mathcal{Y}'_t$ for any $\hat{X} \in \mathcal{L}(\mathcal{H}_S)$ and $\tau \geq t$, though, $\hat{X}_\tau$, $\tau < t$ does not; see Fig. 5.

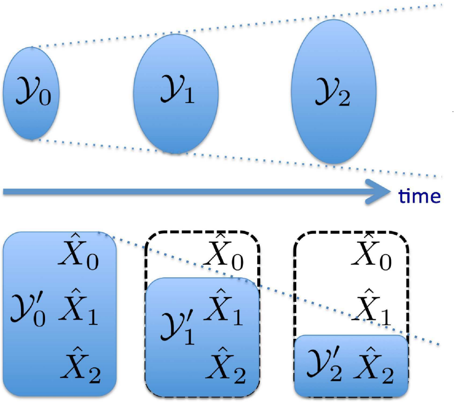

Figure 5. $\mathcal{Y}_t$ and its commutant $\mathcal{Y}'_t$

§ 4.2. Quantum filtering

Let us define $\hat{\pi}_t(\hat{X}) := \mathbb{P}_{\hat{\rho}}[\hat{X}_t \mid \mathcal{Y}_t]$. A formal derivation of the quantum filtering equation is as follows; first, the Doob–Meyer decomposition for the conditional process gives

$$d\hat{\pi}_t(\hat{X}) = \mathbb{P}_{\hat{\rho}}\left[d\hat{\pi}_t(\hat{X})|\ \mathcal{Y}_t\right] + \left(d\hat{\pi}_t(\hat{X}) - \mathbb{P}_{\hat{\rho}}\left[d\hat{\pi}_t(\hat{X})|\ \mathcal{Y}_t\right]\right). \tag{4.8}$$

The first term of Eq. (4.8) implies the prediction from the data up to t and is obtained from the tower property;

$$\begin{aligned}\mathbb{P}_{\hat{\rho}}\left[d\hat{\pi}_t(\hat{X})|\ \mathcal{Y}_t\right] =& \mathbb{P}_{\hat{\rho}}\left[d\hat{X}_t|\ \mathcal{Y}_t\right]\\ =& \hat{\pi}_t\Big(\mathrm{i}[\hat{H},\hat{X}]_-\Big)dt + \frac{1}{2}\hat{\pi}_t\left(\hat{L}^*[\hat{X},\hat{L}]_- + [\hat{L}^*,\hat{X}]_-\hat{L}\right)dt.\end{aligned}$$

The second term of Eq. (4.8) plays the role of the prediction error correction based on the information update, and is martingale. Secondly, we apply the Fujisaki–Kallianpur–Kunita theorem to the second term of Eq. (4.8). Then there exists $\hat{\Xi}_t \in \mathcal{Y}_t$ satisfying

$$\left(d\hat{\pi}_t(\hat{X}) - \mathbb{P}_{\hat{\rho}}\left[d\hat{\pi}_t(\hat{X})|\ \mathcal{Y}_t\right]\right) = \hat{\Xi}_t\left(d\hat{Y}_t - \mathbb{P}_{\hat{\rho}}\left[d\hat{Y}_t|\ \mathcal{Y}_t\right]\right) = \hat{\Xi}_t\left(d\hat{Y}_t - \hat{\pi}_t\left(\hat{L}+\hat{L}^*\right)dt\right).$$

It is possible to determine the $\hat{\Xi}_t \in \mathcal{Y}_t$ from calculating $\mathbb{P}_{\hat{\rho}}[d(\hat{X}_t\hat{Y}_t)\hat{Z}] = \mathbb{P}_{\hat{\rho}}[d(\hat{\pi}_t(\hat{X})\hat{Y}_t)\hat{Z}]$, for all $t \geq 0$ and $\hat{Z} \in \mathcal{Y}_t$. Finally, the quantum filtering equation is given by following equation.

$$\begin{aligned} d\hat{\pi}_t(\hat{X}) =& \hat{\pi}_t\Big(\mathrm{i}[\hat{H},\hat{X}]_-\Big)dt + \frac{1}{2}\hat{\pi}_t\Big(\hat{L}^*[\hat{X},\hat{L}]_- + [\hat{L}^*,\hat{X}]_-\hat{L}\Big)dt \\ &+ \hat{\pi}_t\Big((\hat{L} - \hat{\pi}_t(\hat{L}))^*\hat{X} + \hat{X}(\hat{L} - \hat{\pi}_t(\hat{L}))\Big)(d\hat{Y}_t - \hat{\pi}_t(\hat{L}+\hat{L}^*)dt). \end{aligned} \tag{4.9}$$

$\mathcal{Y}_t$ is identified to a set of classical random variables of the classical probability space $(\Omega, \mathcal{F}, \mathrm{P})$, there exists $\hat{\rho}_t(\omega) \in \mathcal{S}(\mathcal{H}_S)$ for all $\omega \in \Omega$ satisfies

$$\hat{\pi}_t(\hat{X})(\omega) = \mathrm{Tr}[\hat{\rho}_t(\omega)\hat{X}], \quad \forall\hat{X} \in \mathcal{L}(\mathcal{H}_S),\ \forall\omega \in \Omega$$

Using a cyclic property of the trace, the stochastic differential equation of $\hat{\rho}_t$, so-called *the stochastic master equation* or *quantum trajectory equation*, is

$$\begin{aligned} d\hat{\rho}_t =& -\mathrm{i}\left[\hat{H},\hat{\rho}_t\right]_- dt + \left(\hat{L}\hat{\rho}_t\hat{L}^* - \frac{1}{2}\hat{L}^*\hat{L}\hat{\rho}_t - \frac{1}{2}\hat{\rho}_t\hat{L}^*\hat{L}\right)dt \\ &+ \left(\hat{L}\hat{\rho}_t + \hat{\rho}_t\hat{L}^* - \mathrm{Tr}\left[(\hat{L}+\hat{L}^*)\hat{\rho}_t\right]\hat{\rho}_t\right)\left(dy_t - \mathrm{Tr}\left[(\hat{L}+\hat{L}^*)\hat{\rho}_t\right]dt\right). \end{aligned} \tag{4.10}$$

§ 5. Quantum smoothing

§ 5.1. The proposal quantum smoothing

In this section, we consider the fixed point smoothing problem. One of the simplest quantum smoothing setting is the target quantum physical quantity does not evolve under the unitary operator. That implies $[\hat{U}_t, \hat{X}]_- = 0$ and this is called *the Braginsky's quantum nondemolition detection condition* [15]. Since this case can reduce to the filtering problem, it is not a essential quantum smoothing problem. We consider more general setup in this paper. Let us derive the recursive expression of the quantum minimum mean square approximation. We consider the problem of the estimation of $\hat{X}_\tau$, which is the solution of Eq. (4.4) at a fixed time $\tau \geq 0$, from measurement data $\mathcal{Y}_t$ up to $t \geq \tau$. Remember that any element of $\mathcal{Y}_t$ can be seemed as a classical random variable, we can also use the martingale method [40] in order to derive the dynamical estimator. As the rigorous mathematical derivation and jargons make us confuse, we give a sketch how to derive the dynamical estimator. We denote $\mathbb{Q}_t^{\pm}(\hat{X}) := \mathbb{Q}_{\hat{\rho}}^{\pm}[\hat{X} \mid \mathcal{Y}_t]$ for $\hat{X} \in \mathcal{L}(\mathcal{H})$. Since the physical quantity does not evolve and $\mathbb{Q}_t^{\pm}(\hat{X}_\tau) \in \mathcal{Y}_t$ for all $t \geq \tau$, the process $\{\mathbb{Q}_t^{\pm}(\hat{X}_\tau)\}_{t\geq\tau}$ is martingale [40]. An increment of any martingale process can be represented by multiplication between the innovation increment and a uniquely determined coefficient derived from measurement records (*the Fujisaki–Kallianpur–Kunita's* *theorem*).

$$dQ_t^{\pm}\left(\hat{X}_\tau\right) = Q_{t+dt}^{\pm}\left(\hat{X}_\tau\right) - Q_t^{\pm}\left(\hat{X}_\tau\right) = \hat{\Gamma}_t^{\pm}\left(d\hat{Y}_t - \hat{\pi}_t\left(\hat{L}+\hat{L}^*\right)dt\right) \tag{5.1}$$

Note that $Q_\tau^+(\hat{X}_\tau) = \hat{\pi}_\tau(\hat{X})$ and $Q_\tau^-(\hat{X}_\tau) = 0$. Then the problem is to determine the coefficient $\hat{\Gamma}_t^{\pm} \in \mathcal{Y}_t$.

Theorem 5.1 ([43]).
Let $Q_\tau^+\left(\hat{X}_\tau\right) = \hat{\pi}_\tau(\hat{X})$ and $Q_\tau^-\left(\hat{X}_\tau\right) = 0$. Then the recursive estimators are described as following equations;

$$\begin{aligned} dQ_t^{\pm}\left(\hat{X}_\tau\right) =& \frac{1}{2}\Big\{Q_t^+\left(\left[\left(\hat{L}_t+\hat{L}_t^*\right),\hat{X}_\tau\right]_\pm\right) + Q_t^-\left(\left[\left(\hat{L}_t+\hat{L}_t^*\right),\hat{X}_\tau\right]_\mp\right) \\ & - 2Q_t^{\pm}\left(\hat{X}_\tau\right)\hat{\pi}_t\left(\hat{L}+\hat{L}^*\right)\Big\}\left(d\hat{Y}_t - \hat{\pi}_t\left(\hat{L}+\hat{L}^*\right)dt\right), \quad \forall t \geq \tau \end{aligned} \tag{5.2}$$

The sum of the solutions $Q_t^+(\hat{X}_\tau) + Q_t^-(\hat{X}_\tau)$ is the minimum mean square approximation of $\hat{X}_\tau$ from the definition. We call Eq. (5.2) *the real(imaginary) quantum smoother.*

For implementation of the quantum smoother (5.2) is not easy because to calculate the (5.2), we have to consider $Q_t^+\left(\left[\left(\hat{L}_t+\hat{L}_t^*\right),\hat{X}_\tau\right]_+\right)$ and $Q_t^-\left(\left[\left(\hat{L}_t+\hat{L}_t^*\right),\hat{X}_\tau\right]_-\right)$. The time evolution equations of these operator are given by the following lemma.

Lemma 5.2.
The time evolution equation of the $Q_t^{\pm}\left(\hat{R}_t^{\pm}\right)$ and $\hat{R}_t^{\pm} := \left[\left(\hat{L}_t+\hat{L}_t^*\right),\hat{X}_\tau\right]_\pm$, is

$$\begin{aligned} dQ_t^{\pm}\left(\hat{R}_t^{\pm}\right) =& Q_t^{\pm}\left(\left[\mathrm{i}\left[\hat{H}_t,\hat{L}_t+\hat{L}_t^*\right]_-,\hat{X}_\tau\right]_\pm\right)dt \\ & + \frac{1}{2}Q_t^{\pm}\left(\left[\left[\hat{L}_t^*,\hat{L}_t\right]_-\hat{L}_t,\hat{X}_\tau\right]_\pm\right)dt + \frac{1}{2}Q_t^{\pm}\left(\left[\hat{L}_t^*\left[\hat{L}_t^*,\hat{L}_t\right]_-,\hat{X}_\tau\right]_\pm\right)dt \\ & + \Big\{\frac{1}{2}Q_t^+\left(\left[\hat{L}_t+\hat{L}_t^*,\hat{R}_t^{\pm}\right]_\pm\right) + \frac{1}{2}Q_t^-\left(\left[\hat{L}_t+\hat{L}_t^*,\hat{R}_t^{\pm}\right]_\mp\right) \\ & + Q_t^{\pm}\left(\left[\left[\hat{L}_t^*,\hat{L}_t\right]_-,\hat{X}_\tau\right]_\pm\right) + \hat{\pi}_t\left(\hat{L}+\hat{L}^*\right)Q_t^{\pm}\left(\hat{R}_t^{\pm}\right)\Big\} \\ & \times\left(d\hat{Y}_t - \hat{\pi}_t\left(\hat{L}_t+\hat{L}_t^*\right)dt\right). \end{aligned} \tag{5.3}$$

Proof. Let $\hat{R}_t^{\pm} := \left[\left(\hat{L}_t+\hat{L}_t^*\right),\hat{X}_\tau\right]_\pm$. Using Ito rule gives the time evolution equation of $\hat{R}_t$;

$$\begin{aligned}
d\hat{R}_t^{\pm} =& \left[\mathrm{i}\left[\hat{H}_t, \hat{L}_t + \hat{L}_t^*\right]_-, \hat{X}_\tau\right]_\pm dt + \frac{1}{2}\left[\left[\hat{L}_t^*, \hat{L}_t\right]_- \hat{L}_t + \hat{L}_t^*\left[\hat{L}_t^*, \hat{L}_t\right]_-, \hat{X}_\tau\right]_\pm dt \\
&+ \left[\left[\hat{L}_t^*, \hat{L}_t\right]_-, \hat{X}_\tau\right]_\pm d\hat{A}_t + \left[\left[\hat{L}_t, \hat{L}_t^*\right]_-, \hat{X}_\tau\right]_\pm d\hat{A}_t^* \\
=& \left[\mathrm{i}\left[\hat{L}_t + \hat{L}_t^*, \hat{H}_t\right]_-, \hat{X}_\tau\right]_\pm dt + \frac{1}{2}\left[\left[\hat{L}_t^*, \hat{L}_t\right]_- \hat{L}_t + \hat{L}_t^*\left[\hat{L}_t^*, \hat{L}_t\right]_-, \hat{X}_\tau\right]_\pm dt \\
&+ \left[\left[\hat{L}_t^*, \hat{L}_t\right]_-, \hat{X}_\tau\right]_\pm \left(d\hat{A}_t - d\hat{A}_t^*\right).
\end{aligned}$$

From Definitions 3.3 and 3.3, the predictable part of $\{\mathcal{Y}_t\}_t$–adapted process $\mathbb{Q}_t^{\pm}(\hat{R}_t^{\pm})$ satisfies

$$\mathbb{P}_{\hat{\rho}}\left[\hat{Z}d\hat{R}_t^{\pm} \pm d\hat{R}_t^{\pm}\hat{Z}\right] = 2\mathbb{P}_{\hat{\rho}}\left[d\mathbb{Q}_t^{\pm}\left(\hat{R}_t^{\pm}\right)\hat{Z}\right], \ \forall\hat{Z} \in \mathcal{Y}_t.$$

Then

$$\begin{aligned}
\mathbb{P}_{\hat{\rho}}\left[d\mathbb{Q}_t^{\pm}\left(\hat{R}_t^{\pm}\right) \mid \mathcal{Y}_t\right] =& \mathbb{Q}_t^{\pm}\left(\left[\mathrm{i}\left[\hat{H}_t, \hat{L}_t + \hat{L}_t^*\right]_-, \hat{X}_\tau\right]_\pm\right) dt \\
&+ \frac{1}{2}\mathbb{Q}_t^{\pm}\left(\left[\left[\hat{L}_t^*, \hat{L}_t\right]_- \hat{L}_t + \hat{L}_t^*\left[\hat{L}_t^*, \hat{L}_t\right]_-, \hat{X}_\tau\right]_\pm\right) dt.
\end{aligned}$$

The rest of $d\mathbb{Q}_t^{\pm}(\hat{R}_t^{\pm})$ is $\{\mathcal{Y}_t\}_t$–martingale part. From Fujisaki–Kallianpur–Kunita theorem, we obtain

$$\begin{aligned}
d\mathbb{Q}_t^{\pm}\left(\hat{R}_t^{\pm}\right) =& \mathbb{Q}_t^{\pm}\left(\left[\mathrm{i}\left[\hat{H}_t, \hat{L}_t + \hat{L}_t^*\right]_-, \hat{X}_\tau\right]_\pm\right) dt + \frac{1}{2}\mathbb{Q}_t^{\pm}\left(\left[\left[\hat{L}_t^*, \hat{L}_t\right]_- \hat{L}_t, \hat{X}_\tau\right]_\pm\right) dt \\
&+ \frac{1}{2}\mathbb{Q}_t^{\pm}\left(\left[\hat{L}_t^*\left[\hat{L}_t^*, \hat{L}_t\right]_-, \hat{X}_\tau\right]_\pm\right) dt + \hat{\Sigma}_t^{\pm}\left(d\hat{Y}_t - \hat{\pi}_t\left(\hat{L} + \hat{L}^*\right)dt\right),
\end{aligned}$$

where $\hat{\Sigma}_t^{\pm}$ is estimated–to–be. The unknown operator $\hat{\Sigma}_t^{\pm}$ is derived by the equality

$$\mathbb{P}_{\hat{\rho}}\left[\hat{Z}d\left(\hat{Y}_t\hat{R}_t^{\pm}\right) \pm d\left(\hat{R}_t^{\pm}\hat{Y}_t\right)\hat{Z}\right] = 2\mathbb{P}_{\hat{\rho}}\left[d\left(\mathbb{Q}_t^{\pm}\left(\hat{R}_t^{\pm}\right)\hat{Y}_t\right)\hat{Z}\right], \ \forall\hat{Z} \in \mathcal{Y}_t.$$

Finally, we obtain

$$\begin{aligned}
\hat{\Sigma}_t^{\pm} =& \frac{1}{2}\mathbb{Q}_t^{+}\left(\left[\hat{L}_t + \hat{L}_t^*, \hat{R}_t^{\pm}\right]_\pm\right) + \frac{1}{2}\mathbb{Q}_t^{-}\left(\left[\hat{L}_t + \hat{L}_t^*, \hat{R}_t^{\pm}\right]_\mp\right) \\
&+ \mathbb{Q}_t^{\pm}\left(\left[\left[\hat{L}_t^*, \hat{L}_t\right]_-, \hat{X}_\tau\right]_\pm\right) + \hat{\pi}_t\left(\hat{L} + \hat{L}^*\right)\mathbb{Q}_t^{\pm}\left(\hat{R}_t^{\pm}\right).
\end{aligned}$$

□

As in the classical filtering and smoothing theory, it is necessary to derive higher dimensional moments equations to calculate exactly and it usually needs infinitely. There are many approximation methods of the filtering and smoothing equations in classical estimation theory, though, there are a few research about approximation of the quantum dynamical estimators [55]. Some examples are in the next section.

§ 5.2. Special examples

In Gammermark's setting [25], we can consider $[\hat{X}_\tau, \hat{\rho}]_- = 0$ for a certain fixed physical quantity $\hat{X}_\tau$. $\mathbb{Q}_t^-(\hat{X}) = 0$, for all $t \geq 0$, then following corollary holds.

Corollary 5.3. *For given $\hat{X}_\tau$ and $\hat{\rho} \in \mathcal{S}(\mathcal{H})$, assume $[\hat{X}_\tau, \hat{\rho}]_- = 0$. Then*

$$\begin{aligned} d\mathbb{Q}_t^+\left(\hat{X}_\tau\right) =& \frac{1}{2}\Big\{\mathbb{Q}_t^+\left(\left[\left(\hat{L}_t + \hat{L}_t^*\right), \hat{X}_\tau\right]_+\right) - 2\mathbb{Q}_t^+\left(\hat{X}_\tau\right)\hat{\pi}_t\left(\hat{L}+\hat{L}^*\right) \\ & - \mathbb{Q}_t^-\left(\left[\left(\hat{L}_t + \hat{L}_t^*\right), \hat{X}_\tau\right]_-\right)\Big\}\left(d\hat{Y}_t - 2\hat{\pi}_t\left(\hat{L}+\hat{L}^*\right)dt\right), \quad \forall t \geq \tau. \end{aligned} \tag{5.4}$$

Next, we consider the Bragynski's quantum nondemolition detection condition. Suppose that the coupling operator in Eq. (4.1) $\hat{L} \in \mathcal{L}(\mathcal{H}_S)$ is normal and $[\hat{L}, \hat{U}_t]_- = 0$, for all $t \geq 0$. In this case, $\hat{\pi}_t(\hat{L}) = \mathbb{Q}_t^+(\hat{L})$. Then, for $\hat{X}_0 = \hat{X}$

$$\begin{aligned} d\mathbb{Q}_t^+\left(\hat{X}\right) =& \frac{1}{2}\Big\{\mathbb{Q}_t^+\left(\left[\left(\hat{L}+\hat{L}^*\right), \hat{X}\right]_+\right) + \mathbb{Q}_t^-\left(\left[\left(\hat{L}+\hat{L}^*\right), \hat{X}\right]_-\right) \\ & - 2\mathbb{Q}_t^+\left(\hat{X}\right)\mathbb{Q}_t^+\left(\hat{L}+\hat{L}^*\right)\Big\}\left(d\hat{Y}_t - \hat{\pi}_t\left(\hat{L}+\hat{L}^*\right)dt\right), \quad \forall t \geq \tau, \end{aligned} \tag{5.5}$$

and

$$\begin{aligned} d\mathbb{Q}_t^-\left(\hat{X}\right) =& \frac{1}{2}\Big\{\mathbb{Q}_t^+\left(\left[\left(\hat{L}+\hat{L}^*\right), \hat{X}\right]_-\right) + \mathbb{Q}_t^-\left(\left[\left(\hat{L}+\hat{L}^*\right), \hat{X}\right]_+\right) \\ & - 2\mathbb{Q}_t^-\left(\hat{X}\right)\mathbb{Q}_t^+\left(\hat{L}+\hat{L}^*\right)\Big\}\left(d\hat{Y}_t - \hat{\pi}_t\left(\hat{L}+\hat{L}^*\right)dt\right), \quad \forall t \geq \tau. \end{aligned} \tag{5.6}$$

These equations depend on operators in $\mathcal{L}(\mathcal{H}_S)$ and the classical representation implies that $\mathbb{Q}_t^\pm$ can be regarded as linear functional on $\mathcal{L}(\mathcal{H}_S)$. There exist $\hat{\rho}_{0|t}^+$ and $\hat{\rho}_{0|t}^-$ which satisfy $\mathbb{Q}_t^\pm(\hat{X}) = \mathrm{Tr}[\hat{\rho}_{0|t}^\pm \hat{X}]$.

Proposition 5.4. *For given $\hat{X}_0 = \hat{X}$, the quantum smoother is described as follows:*

$$\begin{aligned} d\hat{\rho}_{0|t}^\pm =& \frac{1}{2}\Big\{\left[\hat{\rho}_{0|t}^+, \left(\hat{L}+\hat{L}^*\right)\right]_\pm + \left[\hat{\rho}_{0|t}^-, \left(\hat{L}+\hat{L}^*\right)\right]_\mp \\ & - 2\mathrm{Tr}[\hat{\rho}_t^+(\hat{L}+\hat{L}^*)]\hat{\rho}_{0|t}^\pm\Big\}\left(dy_t - \mathrm{Tr}[\hat{\rho}_t(\hat{L}+\hat{L}^*)]dt\right) \end{aligned} \tag{5.7}$$

where $\hat{\rho}_t$ is a solution of the stochastic master equation (4.10).

§ 6. Conclusion

We introduced the quantum filtering theory and a new quantum smoothing theory. The orthogonality is the key word of this paper and it plays an important role in the derivation of the recursive estimators. The future work is how to implement of the smoother in practice.

References

[1] Y. Aharonov, D. Z. Albert, and L. Vaidman. How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100. *Physical Review Letters*, 60(14):1351–1354, 1988.

[2] Y. Aharonov, P. G. Bergmann, and J. L. Lebowitz. Time symmetry in the quantum process of measurement. *Physical Review*, 134(6B):B1410–B1416, 1964.

[3] C. Altafini and F. Ticozzi. Modeling and Control of Quantum Systems: An Introduction. *IEEE Transaction on Automatic Control*, 57(8):1898–1917, 2012.

[4] S. Amari and H. Nagaoka. *Methods of Information Geometry.* Amer Mathematical Society, 2007.

[5] H. Amini, Z. Miao, Y. Pan, M. R. James, and H. Mabuchi. On the generalization of linear least mean squares estimation to quantum systems with non-commutative outputs. *e-print arXiv:1406.4599*, 2014.

[6] A. Bain and D. Crişan. *Fundamentals of Stochastic Filtering*, . Springer Verlag, 2009.

[7] K. Banaszek and K. Wódkiewicz. Operational theory of homodyne detection. *Physical Review A*, 55:3117–3123, Apr 1997.

[8] V. P. Belavkin. Nondemolition Principle of Quantum Measurement Theory. *Foundations of Physics (Historical Archive)*, 24(5):685–714, 1994.

[9] V. P. Belavkin. Eventum mechanics of quantum trajectories: Continual measurements, quantum predictions and feedback control. *e-print math-ph/0702079*, 2007.

[10] V. P. Belavkin and Madalin Guţă, editors. *Quantum Stochastics and Information: Statistics, Filtering and Control.* World Scientific, 2008.

[11] P. Billingsley. *Probability and Measures.* Wiley & Sons, Anniversary edition, 2012.

[12] Jiri Blank, Pavel Exner, and Miloslav Havlicek. *Hilbert Space Operators in Quantum Physics.* Springer, 2 edition, 2008.

[13] L. Bouten, R. van Handel, and M. R. James. An Introduction to Quantum Filtering. *SIAM Journal on Control and Optimization*, 46(6):2199–2241, 2007.

[14] L. Bouten, R. van Handel, and M. R. James. A Discrete Invitation to Quantum Filtering and Feedback Control. *SIAM Review*, 51(2):239–316, 2009.

[15] V. B. Braginsky, Y. I. Vorontsov, and K. S. Thorne. Quantum Nondemolition Measurements. *Science*, 209:547–557, 1980.

[16] H. P. Breuer and F. Petruccione. *The Theory of Open Quantum Systems.* Clarendon Press, Oxford, 2006.

[17] R. S. Bucy and P. D. Joseph. *Filtering for Stochastic Processes with Applications to Guidance.* Chelsea Pub Co, new edition, 2005.

[18] J. Combes, C. Ferrie, Z. Jiang, and C. M. Caves. Quantum limits on postselected, probabilistic quantum metrology. *Physical Review A*, 89:052117, May 2014.

[19] D. Das and Arvind. Estimation of quantum states by weak and projective measurements. *Physical Review A*, 89:062121, Jun 2014.

[20] F. de los Santos, C. Guglieri, and E. Romera. Application of new Rényi uncertainty relations to wave packet revivals. *Physica E: Low-dimensional Systems and Nanostructures*, 42(3):303–307, 2010.

[21] D. Dong and I. R. Petersen. Quantum control theory and applications : A survey. *IET Control Theory & Applications*, 4(12):2651–2671, 2010.

[22] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd. Understanding quantum weak values: Basics and applications. *Reviews of Modern Physics*, 86:307–316, Mar 2014.

[23] T. E. Duncan. Some filtering results in Riemann manifolds. *Information and Control*, 35(3):182–195, 1977.

[24] C. Ferrie and J. Combes. Weak Value Amplification is Suboptimal for Estimation and Detection. *Physical Review Letters*, 112:040406, Jan 2014.

[25] S. Gammelmark, B. Julsgaard, and K. Mølmer. Past Quantum States of a Monitored System. *Physical Review Letters*, 111:160401, Oct 2013.

[26] C. W. Gardiner and P. Zoller. *Quantum Noise : A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods With Applications to Quantum Optics*. Springer–Verlag, 3 edition, 2004.

[27] A. Guichardet. *Tensor products of C^*-algebras*. Århus Universitet, Matematisk Institut, 1969.

[28] D. Guo, S. Shamai, and S. Verdú. Mutual information and minimum mean-square error in Gaussian channels. *IEEE Transactions on Information Theory*, 51(4):1261–1282, 2005.

[29] D. Guo, Y. Wu, S. Shamai, and S. Verdú. Estimation in Gaussian noise: Properties of the minimum mean-square error. *IEEE Transactions on Information Theory*, 57(4):2371–2385, 2011.

[30] S. Haykin. *Adaptive Filter Theory*. Prentice hall, 4 edition, 2001.

[31] C. W. Helstrom. Minimum mean-squared error of estimates in quantum statistics. *Physics Letters A*, 25(2):101–102, 1967.

[32] C. W. Helstrom. *Quantum Detection and Estimation Theory*. Academic Press, 1976.

[33] A. S. Holevo. *Probability and Statistical Aspects of Quantum Theory*. New York: North-Holland, 1982.

[34] A. S. Holevo. *Statistical Structure of Quantum Theory*. Springer, 2001.

[35] A. S. Holevo. *Quantum Systems, Channels, Information: A Mathematical Introduction*. De Gruyter, 2012.

[36] R. L. Hudson and K. R. Parthasarathy. Quantum Ito's Formula and Stochastic Evolutions. *Communications in Mathematical Physics*, 93(3):301–323, 1984.

[37] M. James. Quantum Control Theory. *e-print arXiv:1406.5260*, 2014.

[38] Gopinath Kallianpur. *Stochastic Filtering Theory*. Springer-Verlag, 1980.

[39] F. Kaneda, S.-Y. Baek, M. Ozawa, and K. Edamatsu. Experimental Test of Error-Disturbance Uncertainty Relations by Weak Measurement. *Physical Review Letters*, 112:020402, Jan 2014.

[40] V. Krishnan. *Nonlinear Filtering and Smoothing : An Introduction to Martingales, Stochastic Integrals and Estimation*. John Wiley & Sons, 1984.

[41] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[42] K. Ohki. A quantum recursive least mean square estimation for past open quantum

dynamical systems. In *SICE Annual Conference*, 2014.

[43] K. Ohki. A smoothing theory for open quantum systems: The least mean square approach. In *54th IEEE Conference on Decision and Control*, pages 4350 – 4355, 2015.

[44] A. Peres. *Quantum Theory: Concepts and Methods.* Kluwer Academic Publishers, 1993.

[45] M. Ringbauer, D. N. Biggerstaff, M. A. Broome, A. Fedrizzi, Cyril Branciard, and Andrew G. White. Experimental Joint Quantum Measurements with Minimum Uncertainty. *Physical Review Letters*, 112:020401, Jan 2014.

[46] E. Romera and F. de los Santos. Fractional revivals through Rényi uncertainty relations. *Physical Review A*, 78:013837, Jul 2008.

[47] E. Romera and F. de los Santos. Identifying Wave-Packet Fractional Revivals by Means of Information Entropy. *Physical Review Letters*, 99:263601, Dec 2007.

[48] E. Romera and F. de los Santos. Fisher information, nonclassicality and quantum revivals. *Physics Letters A*, 377(37):2284–2287, 2013.

[49] A. H. Sayed. *Adaptive Filters.* John Wiley & Sons, 2008.

[50] D. Simon. *Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches.* John Wiley & Sons, 2006.

[51] S. M. Stigler. *The History of Statistics: The Measurement of Uncertainty Before 1900.* Harvard University Press, 1986.

[52] M. Takesaki. *Theory of Operator Algebras 1.* Springer, 2003.

[53] M. Tsang. Time-symmetric quantum theory of smoothing. *Physical Review Letters*, 102(25):250403, 2009.

[54] M. Tsang. A Bayesian quasi-probability approach to inferring the past of quantum observables. *e-print arXiv:1403.3353*, 2014.

[55] R. van Handel and Hideo Mabuchi. Quantum projection filter for a highly nonlinear model in cavity QED. *Journal of Optics B: Quantum and Semiclassical Optics*, 7:S226, 2005.

[56] H. L. Van Trees, K. L. Bell, and Z. Tian. *Detection, Estimation, and Modulation Theory: Part 1 Detection, Estimation, and Filtering Theory.* John Wiley & Sons, 2 edition, 2013.

[57] S. Verdú. Mismatched estimation and relative entropy. *IEEE Transactions on Information Theory*, 56(8):3712–3720, 2010.

[58] D. F. Walls and G. J. Milburn. *Quantum Optics.* Springer Verlag, 2 edition, 2008.

[59] Y. Watanabe, T. Sagawa, and M. Ueda. Uncertainty relation revisited from quantum estimation theory. *Physical Review A*, 84:042121, Oct 2011.

[60] N. Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series.* Wiley New York, 1950.

[61] H. M. Wiseman and G. J. Milburn. *Quantum Measurement and Control.* Cambridge University Press, 2009.

[62] M. Yanagisawa. Quantum smoothing. *e-print arXiv:0711.3885*, 2007.

[63] H. Yonezawa, D. Nakane, T. A. Wheatley, K. Iwasawa, S. Takeda, H. Arao, K. Ohki, K Tsumura, D.W. Berry, T. C. Ralph, H. M. Wiseman, E. H. Huntington, and A. Furusawa. Quantum-Enhanced Optical-Phase Tracking. *Science*, 337(6101):1514–1517, 2012.

[64] J. Zhang, Y.-x. Liu, R.-B. Wu, K. Jacobs, and F. Nori. Quantum feedback: theory, experiments, and applications. *e-print arXiv:1407.8536*, 2014.